Lecture No. 7

Introduction To Statistics, Statistics And Probability

Dr. Shabbir Ahmad

Assistant Professor,
Department of Mathematics,
COMSATS University
Islamabad, Wah Campus

Measures of Dispersion

Different Tools (Q.D, CQD, S.D, Variance, C.V)

Dr. Shabbir Ahmad
Assistant Professor, Department of Mathematics, COMSATS University Islamabad, Wah Campus
Cell # 0323-5332733, 0332-5332733. Date: 10/4/2021 10:52:42 PM

In this lecture

- Quartiles Deviation?
- Coefficient of Quartile Deviation?
- Merits and Demerits of Quartile Deviation
- Variance and Standard Deviation
- Coefficient of Variation

Dr. Shabir Ahmad
Assistant Professor, Department of Mathematics, COMSATS University Islamabad, Wah Campus
Cell # 0323-5332733, 0332-5332733. Date: 10/4/2021 10:52:42 PM

Quartiles Deviation:

- *Quartile deviation (Q.D)* means the semi variation between the upper quartile (Q_3) and the lower quartile (Q_1) in a distribution. *OR*
- The *quartiles deviation* is defineD as half the difference between third and first quartiles, and is defined as:

$$Q.D = \frac{Q_3 - Q_1}{2}$$

- The difference between upper quartile (Q_3) and the lower quartile (Q_1) is called *semi inter quartiles range*, is the spread of 50% of the data values. SO,

$$\text{IQR} = Q_3 - Q_1$$

Dr. Shabbir Ahmad
Assistant Professor, Department of Mathematics, COMSATS University Islamabad, Wah Campus
Cell # 0323-5332733, 0332-5332733. Date: 10/4/2021 10:53:39 PM

Coefficient of Quartile Deviation:

- A relative measure of dispersion based on the quartile deviation is called coefficient of quartile deviation.
- We calculate the Coefficient of $Q.D$ as;

$$\text{Coefficient of } Q.D = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

Dr. Shabbir Ahmad
Assistant Professor, Department of Mathematics, COMSATS University Islamabad, Wah Campus
Cell # 0323-5332733, 0332-5332733. Date: 10/4/2021 10:53:58 PM

Q.D for Ungrouped Data:

- ***First:*** Arrange the data in ascending order
- ***Second:*** Find lower (first) quartile.
- ***Third:*** Find upper (third) quartile.
- ***Finally:*** Put the values in the formula of Quartile deviation.

Upper and Lower Quartiles are calculated by using formula:

$$Q_i = \frac{i(n+1)}{4} th\ observation$$

Where *i=1 and 3*

Example 1:

Example 1: Find quartiles deviation and coefficient of quartiles deviation.

3,7,12,25,37,48,15,69,52,73,70,88,82,80,92.

Solution: First we arranging data into ascending order;

3,7,12,15,25,37,48,52,69,70,73,80,82,88,92.

$$Q_1\ No. = \frac{1(n+1)}{4}$$

$$Q_1\ No. = \frac{1(15+1)}{4} = \frac{16}{4}$$

$$Q_1\ No. = 4$$

Its means that the four number value of arranged data is first quartiles

$$\boldsymbol{Q_1 = 15}$$

Dr. Shabbir Ahmad
Assistant Professor, Department of Mathematics, COMSATS University Islamabad, Wah Campus
Cell # 0323-5332733, 0332-5332733. Date: 10/4/2021 10:54:06 PM

Example 1 (Cont.):

$$Q_3\ No. = \frac{3(n+1)}{4}$$

$$Q_3\ No. = \frac{3(15+1)}{4} = \frac{48}{4} = 12$$

Its means that the twelve number value of arranged data is third quartiles

$$\boldsymbol{Q_3 = 80}$$

We calculate Absolute dispersion the $Q.D$ as;

$$Q.D = \frac{Q_3 - Q_1}{2}$$

$$Q.D = \frac{80 - 15}{2}$$

$$\boldsymbol{Q.D = 32.5}$$

Dr. Shabbir Ahmad
Assistant Professor, Department of Mathematics, COMSATS University Islamabad, Wah Campus
Cell # 0323-5332733, 0332-5332733. Date: 10/4/2021 10:54:01 PM

Example 1 (Cont.):

We calculate the relative dispersion, Coefficient of $Q.D$ as;

$$\text{Coefficient of } Q.D = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

$$\text{Coefficient of } Q.D = \frac{80 - 15}{80 + 15}$$

$$\text{Coefficient of } Q.D = \frac{65}{95}$$

$$\textbf{Coefficient of } \boldsymbol{Q.D = 0.6842}$$

Dr. Shabbir Ahmad
Assistant Professor, Department of Mathematics, COMSATS University Islamabad, Wah Campus
Cell # 0323-5332733, 0332-5332733. Date: 10/4/2021 10:54:09 PM

Example 2:

Example 2: Find Quartiles Deviation and Coefficient of Quartile Deviation for discrete data

X	0	1	2	3	4	5	6	7	8
f	1	9	26	59	72	52	29	7	1

Solution: Now we calculate only cumulative frequency

X	*f*	*Cmf (cumulate frequency)*
0	1	1
1	9	10
2	26	36
3	59	95
4	72	167
5	52	219
6	29	248
7	7	255
8	1	256
	$\sum f = 256$	

$$Q_1\ No. = \frac{1(n+1)}{4}$$

Here $n = \sum f$

$$Q_1\ No. = \frac{1(256+1)}{4} = \frac{257}{4} = 64.25$$

$$\boldsymbol{Q_1 = 3}$$

$$Q_3\ No. = \frac{3(n+1)}{4}$$

$$Q_3\ No. = \frac{3(256+1)}{4} = \frac{771}{4} = 192.75$$

$$\boldsymbol{Q_3 = 5}$$

Dr. Shabbir Ahmad
Assistant Professor, Department of Mathematics, COMSATS University Islamabad, Wah Campus
Cell # 0323-5332733, 0332-5332733. Date: 10/4/2021 10:54:18 PM

Example 2 (Cont.):

We calculate Absolute dispersion the $Q.D$ as;

$$Q.D = \frac{Q_3 - Q_1}{2}$$

$$Q.D = \frac{5 - 3}{2}$$

$$\boldsymbol{Q.D = 1}$$

We calculate the relative dispersion, Coefficient of $Q.D$ as;

$$\text{Coefficient of } Q.D = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

$$\text{Coefficient of } Q.D = \frac{5 - 3}{5 + 3} = \frac{2}{8}$$

$$\textbf{Coefficient of } \boldsymbol{Q.D = 0.25}$$

Dr. Shabbir Ahmad
Assistant Professor, Department of Mathematics, COMSATS University Islamabad, Wah Campus
Cell # 0323-5332733, 0332-5332733. Date: 10/4/2021 10:54:22 PM

Merits and Demerits of Quartile Deviation:

Merits:

1. It is easy to understand and easy to calculate.
2. It is least affected by extreme values.
3. It can be used in open-ended frequency distribution.

Demerits:

1. It is not suited to algebraic treatment.
2. It is very much affected by sampling fluctuations.
3. The method of dispersion is not based on all the items of series.
4. It ignores 50% of the distribution.

Variance and Standard Deviation:

- The variance of the observation is defined as the sum of the square of deviation of the values from their mean then divided by n, and it is denoted by $Var(X)$ (say variance of X)
- For sample data variance is denoted by S^2, and if the variance is calculated for a population is denoted by σ^2.
- The standard deviation is define as the positive square root of the sum of the squared deviation of the values from their mean then divided by n, and it is denoted by $S.D(X)$.
- For sample data standard deviation is denoted by S, and if the standard deviation is calculated for a population is denoted by σ.
- The formula for variance and standard deviation (for ungroup data) of a set of n values, $X_1, X_2, X_3, \dots, X_n$, is defined as:

$$Variance = \frac{\sum(X - \overline{X})^2}{n}$$

$$S.D = \sqrt{\frac{\sum(X - \overline{X})^2}{n}}$$

Dr. Shabbir Ahmad
Assistant Professor, Department of Mathematics, COMSATS University Islamabad, Wah Campus
Cell # 0323-5332733, 0332-5332733. Date: 10/4/2021 10:54:30 PM

Variance and Standard Deviation:

The formula used to calculate standard deviation and variance of Discrete and Continuous frequency distribution.

$$Variance = \frac{\sum f(X-\overline{X})^2}{\sum f} \; AND \; S.D = \sqrt{\frac{\sum f(X-\overline{X})^2}{\sum f}}$$

for group data.

The relative measure of variance and standard deviation is called Coefficient of variation, and is calculated by using formula:

$$\text{Coefficient of variance} = \frac{S.D}{\overline{X}} \times 100$$

This can be used to compare two distributions directly to see which has more dispersion because it does not depend on units of the distribution.

Dr. Shabbir Ahmad
Assistant Professor, Department of Mathematics, COMSATS University Islamabad, Wah Campus
Cell # 0323-5332733, 0332-5332733. Date: 10/4/2021 10:54:26 PM

Example 3:

Example 3. Find S.D and Variance for the following data 2,4,5,8,6.

Solution.

X	$(X-\overline{X})$ $=(X-5)$	$(X-\overline{X})^2$
2	-3	9
4	-1	1
5	0	0
8	3	9
6	1	1
$\sum X = 25$		$\sum(X-\overline{X})^2 = 20$

$$\overline{X} = \frac{\sum X}{n} = \frac{25}{5}$$

$$\overline{X} = 5$$

$$S.D = \sqrt{\frac{\sum(X-\overline{X})^2}{n}}$$

$$S.D = \sqrt{\frac{20}{5}}$$

$$\boldsymbol{S.D = 2}$$

Variance is square of S.D,

$$Variance = (S.D)^2 = (2)^2$$

$$\boldsymbol{Variance = 4}$$

Dr. Shabbir Ahmad
Assistant Professor, Department of Mathematics, COMSATS University Islamabad, Wah Campus
Cell # 0323-5332733, 0332-5332733. Date: 10/4/2021 10:55:12 PM

Coefficient of Variation:

Coefficient of Variance is,

$$\text{Coefficient of variance} = \frac{S.D}{\overline{X}} \times 100$$

$$\text{Coefficient of variance} = \frac{2}{5} \times 100$$

$$\mathbf{Coefficient\ of\ variance = 40\%}$$

Dr. Shabbir Ahmad
Assistant Professor, Department of Mathematics, COMSATS University Islamabad, Wah Campus
Cell # 0323-5332733, 0332-5332733. Date: 10/4/2021 10:55:15 PM

Example 4:

Example 4. Find Standard deviation, Variance also find coefficient of variance**.**

classes	20-24	25-29	30-34	35-39	40-44	45-49	50-54
f	1	4	8	11	15	9	2

Solution.

classes	*f*	*X*	*fX*	$(X-\overline{X})$	$(X-\overline{X})^2$	$f(X-\overline{X})^2$
20-24	1	22	22	-17	289	289
25-29	4	27	108	-12	144	576
30-34	8	32	256	-7	49	392
35-39	11	37	407	-2	4	44
40-44	15	42	630	3	9	135
45-49	9	47	423	8	64	576
50-54	2	52	104	13	169	338
	$\sum f = 50$		$\sum fX = 1950$			$\sum f(X-\overline{X})^2 = 2350$

$$\overline{X} = \frac{\sum fX}{\sum f} = \frac{1950}{50}$$

$$\overline{X} = 39$$

Dr. Shabbir Ahmad
Assistant Professor, Department of Mathematics, COMSATS University Islamabad, Wah Campus
Cell # 0323-5332733, 0332-5332733. Date: 10/4/2021 10:55:18 PM

Example 4 (Cont.):

$$S.D = \sqrt{\frac{\sum f\left(X - \overline{X}\right)^2}{\sum f}}$$

$$S.D = \sqrt{\frac{2350}{50}}$$

$$\boldsymbol{S.D = 6.86}$$

Variance is square root of S.D,

$$Variance = (S.D)^2$$

$$Variance = (6.86)^2$$

$$\boldsymbol{Variance = 47}$$

Dr. Shabbir Ahmad
Assistant Professor, Department of Mathematics, COMSATS University Islamabad, Wah Campus
Cell # 0323-5332733, 0332-5332733. Date: 10/4/2021 10:55:20 PM

Example 4 (Cont.)

Coefficient of Variance is,

$$\text{Coefficient of variance} = \frac{S.D}{\overline{X}} \times 100$$

$$\text{Coefficient of variance} = \frac{6.86}{39} \times 100$$

$$\mathbf{Coefficient\ of\ variance = 17.59\%}$$

Dr. Shabbir Ahmad
Assistant Professor, Department of Mathematics, COMSATS University Islamabad, Wah Campus
Cell # 0323-5332733, 0332-5332733. Date: 10/4/2021 10:55:24 PM

Comparing Coefficients of Variation

- Stock A:
  - *Average price last year = $50*
  - *Standard deviation = $5*

$$CV_A = \left(\frac{S}{\overline{X}} \right) \cdot 100\% = \frac{\$5}{\$50} \cdot 100\% = 10\%$$

- Stock B:
  - *Average price last year = $100*
  - *Standard deviation = $5*

$$CV_B = \left(\frac{S}{\overline{X}} \right) \cdot 100\% = \frac{\$5}{\$100} \cdot 100\% = 5\%$$

Both stocks have the same standard deviation, but stock B is less variable relative to its price

Dr. Shabbir Ahmad
Assistant Professor, Department of Mathematics, COMSATS University Islamabad, Wah Campus
Cell # 0323-5332733, 0332-5332733. Date: 10/4/2021 10:55:26 PM

20

Questions for Practice:

Question 1: The following sample data of the number of communications are taken from logs of communication with Distance Education students: 5, 9, 5, 23, 27, 55, 34, 7, 30, 15, 22, 60, 14, 52, 297, 8, 51, 15, 51, 35, 15, 39, 137, 43, 38, 14, 93, 7. Calculate Q.D, Coefficient of Q.D, Variance ,Standard Deviation and Coefficient of Variation.

Question 2: The following is the age distribution of 1000 persons working in an organization:

Age Group	Number of Persons
20-25	30
25-30	160
30-35	210
35-40	180
40-45	145
45-50	105
50-55	70
55-60	60
60-65	40

Calculate Q.D, Coefficient of Q.D, Variance, Standard Deviation and Coefficient of Variation..

Dr. Shabbir Ahmad
Assistant Professor, Department of Mathematics, COMSATS University Islamabad, Wah Campus
Cell # 0323-5332733, 0332-5332733. Date: 10/4/2021 10:55:37 PM

ANY QUESTION