Lecture 4:

16 February 2023 18:17

3 Mean
mode
Median

Measures of Variation

rarely

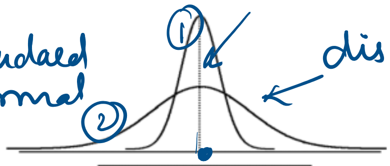
4

- Variation
  - Range
  - Variance
  - Standard Deviation
  - Coefficient of Variation

dataset

- Measures of variation give information on the spread or variability or dispersion of the data values.

Standard normal



distribution of data

Same center, different variation

Mean = Median = mode

Measures of Variation: The Range

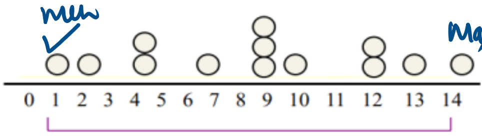
- Simplest measure of variation
- Difference between the largest and the smallest values:

$$\text{Range} = X_{\text{largest}} - X_{\text{smallest}}$$

Example:

min

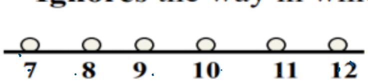
Max



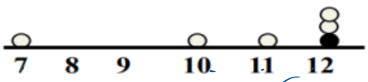
Range = 14 - 1 = 13

Measures of Variation: Why The Range Can Be Misleading

- **Ignores** the way in which data are **distributed**



Range = 12 - 7 = 5



Range = 12 - 7 = 5

- Sensitive to outliers

1,1,1,1,1,1,1,1,1,1,1,2,2,2,2,2,2,2,2,3,3,3,3,4,5

Range = 5 - 1 = 4

1,1,1,1,1,1,1,1,1,1,1,2,2,2,2,2,2,2,2,3,3,3,3,4,120

120 → 0

Range = 120 - 1 = 119

Interfractile range: Median is 0.5 fractile.

First third	Second third	Last third
863	1698	1138
903	1745	1204
957	1802	1354
1041	1883	1624

1/3

3/3

Interfractile range: Median is 0.5 fractile.

First third	Second third	Last third
863		1138
	1698	
903		1204
	1745	
957		1354
	1802	
1041		1624
	1883	

1/3 | 3/3

1/3 fractile | 2/3 fractile

10 parts | 4 parts | 100 parts

If we divide data in??????? deciles , quartile and percentile

$\boxed{\textbf{Interquartile range: Q3-Q1}}$

2

Measures of Variation: The Variance

- Average (approximately) of squared deviations of values from the mean

$\bar{X} : x_i$

– Sample variance:
$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$$

Population variance → $\frac{}{n}$

Where $\bar{X}$ = arithmetic mean

n = sample size

X_i = i^{th} value of the variable X

3

Measures of Variation: The Standard Deviation

- **Most commonly** used measure of variation
- Shows variation about the **mean**
- Is the **square root of the variance**
- Has the same units as the original data

Mean

– Sample standard deviation:
$$S = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}}$$

Measures of Variation: The Standard Deviation

Steps for Computing Standard Deviation

1. Compute the **difference between each value** and the **mean**.
2. **Square** each difference.
3. **Add the squared** differences.
4. **Divide this total by n-1 to** get the sample variance.
5. Take the **square root of the sample** variance to get the sample standard deviation.

$$\sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}}$$

Measures of Variation: Sample Standard Deviation

Example

ungrouped

Sample Data (X_i) : **10 12 14 15 17 18 18 24**

n = 8 Mean = $\overline{X}$ = 16

$$S = \sqrt{\frac{(10-\overline{X})^2 + (12-\overline{X})^2 + (14-\overline{X})^2 + \cdots + (24-\overline{X})^2}{n-1}}$$

$$= \sqrt{\frac{(10-16)^2 + (12-16)^2 + (14-16)^2 + \cdots + (24-16)^2}{8-1}}$$

$$= \sqrt{\frac{130}{7}} = 4.3095$$

S.D -

Sample variance

Mean=1.351

Observation (x) (1)	Mean $(\bar{x})$ (2)	$x - \bar{x}$ (1) − (2)	$(x - \bar{x})^2$ [(1) − (2)]^2	x^2 (1)^2
863	1,351	−488	238,144	744,769
903	1,351	−448	200,704	815,409
957	1,351	−394	155,236	915,849
1,041	1,351	−310	96,100	1,083,681
1,138	1,351	−213	45,369	1,295,044
1,204	1,351	−147	21,609	1,449,616
1,354	1,351	3	9	1,833,316
1,624	1,351	273	74,529	2,637,376
1,698	1,351	347	120,409	2,883,204
1,745	1,351	394	155,236	3,045,025
1,802	1,351	451	203,401	3,247,204
1,883	1,351	532	283,024	3,545,689
			$\Sigma(x - \bar{x})^2 \rightarrow$ 1,593,770	23,496,182 $\leftarrow \Sigma x^2$

Variance:)

$$\boxed{s^2 = \frac{\Sigma(x-\bar{x})^2}{n-1}}$$

$$= \frac{1{,}593{,}770}{11}$$

= 144,888 (or 144,888 [thousands of dollars]^2) ← Sample variance

$$s = \sqrt{s^2}$$

$$= \sqrt{144{,}888}$$

= 380.64 (that is, $380,640) ← Sample standard deviation

Standard Deviation (Sample) for Grouped Data

Ex :- Frequency Distribution of Return on Investment of Mutual Funds

Return on Investment	Number of Mutual Funds
5-10	10
10-15	12

Standard Deviation (Sample) for Grouped Data

Ex :- Frequency Distribution of Return on Investment of Mutual Funds

Return on Investment	Number of Mutual Funds
5-10	10
10-15	12
15-20	16
20-25	14
25-30	8
Total	**60**

Solution for the Example

$$\frac{\Sigma f(x-\bar{x})^2}{n-1}$$

$\frac{5}{2} = 2.5$

A	B	C	D	E	F	G	H
1	Return on	Investment		No of			
2	class		MidPoint	Funds	fxx		
3	Lower limit	Upper Limit	X	f	fx	$(x-\bar{x})^2$	$f(x-\bar{x})^2$
4	5	10	7.50	10	75	96.69	966.94
5	10	15	12.50	12	150	23.36	280.33
6	15	20	17.50	16	280	0.03	0.44
7	20	25	22.50	14	315	26.69	373.72
8	25	30	27.50	8	220	103.36	826.89
9				Σ= 60	1040	Σ=	2448.33
10			$\bar{x}$	Mean=	17.333		
11				Sample Variance =			41.50
12			√	Sample Standard Deviation =			6.44

$$\frac{\Sigma f \times x}{n}$$

Class	Frequency
700-799	4
800-899	7
900	8
1000	10
1100	12
1200	17
1300	13
1400	10
1500	9
1600	7
1700	2
1800-1899	1

$n-1 \Rightarrow n.$

Find population's standard deviation S.D.

Class	Midpoint x (1)	Frequency f (2)	$f \times x$ (3) = (2) × (1)	Mean μ (4)	$x - \mu$ (1) − (4)	$(x-\mu)^2$ $[(1)-(4)]^2$	$f(x-\mu)^2$ (2) × $[(1)-(4)]^2$
700– 799	750	4	3,000	1,250	−500	250,000	1,000,000
800– 899	850	7	5,950	1,250	−400	160,000	1,120,000
900– 999	950	8	7,600	1,250	−300	90,000	720,000
1,000–1,099	1,050	10	10,500	1,250	−200	40,000	400,000
1,100–1,199	1,150	12	13,800	1,250	−100	10,000	120,000
1,200–1,299	1,250	17	21,250	1,250	0	0	0
1,300–1,399	1,350	13	17,550	1,250	100	10,000	130,000
1,400–1,499	1,450	10	14,500	1,250	200	40,000	400,000
1,500–1,599	1,550	9	13,950	1,250	300	90,000	810,000
1,600–1,699	1,650	7	11,550	1,250	400	160,000	1,120,000
1,700–1,799	1,750	2	3,500	1,250	500	250,000	500,000
1,800–1,899	1,850	1	1,850	1,250	600	360,000	360,000
		Σ= 100	Σ= 125,000				Σ⇒ 6,680,000

$$\bar{x} = \frac{\Sigma(f \times x)}{n} = \frac{125,000}{100}$$

= 1,250 (thousands of dollars) ← Mean

[3-3]

$$\bar{x} = \frac{\Sigma(f \times x)}{n} \qquad [3\text{-}3]$$

$$= \frac{125{,}000}{100}$$

$$= 1{,}250 \text{ (thousands of dollars)} \leftarrow \text{Mean}$$

$$\sigma^2 = \frac{\Sigma f(x - \mu)^2}{N} \qquad [3\text{-}15]$$

$$= \frac{6{,}680{,}000}{100}$$

Variance

$$= 66{,}800 \text{ (or } 66{,}800 \text{ [thousands of dollars]}^2) \leftarrow \text{Variance}$$

$$\sigma = \sqrt{\sigma^2} \qquad [3\text{-}16]$$

$$= \sqrt{66{,}800}$$

$$= 258.5 \leftarrow \text{Standard deviation} = \$258{,}500$$

---

4)

Measures of Variation: Comparing Standard Deviations

The **coefficient of variation** (CV) is a measure of relative **variability**.

It is the ratio of the **standard deviation to the mean** (average).

$\frac{SD}{x}$.

Always in percentage (%)

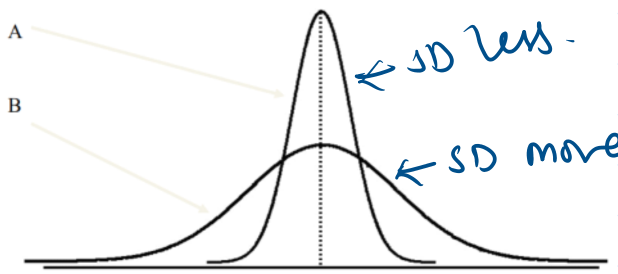
Shows **variation relative to mean**

Can be used to compare the variability of two or more sets of data measured in **different units**

$$\boxed{CV = \left(\frac{S}{\bar{X}}\right) \cdot 100\%}$$

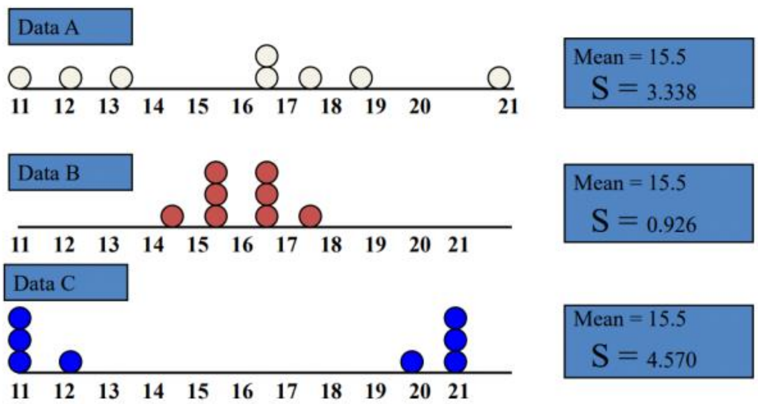
Measures of Variation: Comparing Standard Deviations

Which curve has higher SD?



Measures of Variation: Comparing Standard Deviations

The **coefficient of variation** (CV) is a measure of relative **variability**. It is the ratio of the **standard deviation to the mean** (average).



$\frac{SD}{x} \times 100$

Measures of Variation:
Comparing Coefficients of Variation

- Stock A:
  - Average price last year = \$50 $\bar{x}$
  - Standard deviation = \$5 s

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

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

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

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

Terminologies ⇒ Population Sample

1) Mean μ $\bar{x}$

$$\mu = \sum_{i=1}^{N} \frac{x_i}{N} \qquad \bar{x} = \sum_{i=1}^{n} \frac{x_i}{n}$$

2) Variance σ^2 s^2

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N} \qquad s^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}$$

3) standard deviation

$$\sigma = \sqrt{\sigma^2} \qquad s = \sqrt{s^2}$$