# Unit 1: Measures of Central Tendancy
# Measures of Dispersion of a Distribution

Definition: a distribution: a set of elements/ data points.

marks (%)
ex: 91, 91, 91, 93, 100, 100
(a quantitative d)

ex. green, pink, black, blue, green Bernie Sanders
ex. Sanders, Sanders, Sanders, Clinton, Sanders
(a qualitative d)
↳ identify mode

# Measures of Central Tendancy

. Mode (Mo): the data point w highest frequency

. Mean / Average:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x_1 + x_2 + x_3 + \cdots + x_n}{n}$$

where to end

Sigma "take the sum of"

i - initial where to start

n - total # of data points in distribution

. median (md): the middle data point of a sorted distribution

ex. 6, 3, 5, 4, 4

3, 4, 4, 5, 6

3, 4, 4, 5, 6, 6

their average is the median = 4.5

position/rank of median

$$\frac{n+1}{2}$$

$$\frac{6+1}{2}$$

$$\frac{7}{2} = 3.5$$

take avg between 3rd and 4th data point

Marks %

66, 66, 66, 66, 66, (66), 78, 88, 92, 94, 95

Mode = 66%
Median = 66%
Avg = 76.6%

~~30~~, 60, 72, 72, 74, 76, 77, 80

4.5

4 5

Median = 73%

Mode: = 72%

Avg: 67.62%

## Measures of Dispersion

↳ they tell us how reliable the measures of central tendancy are

↳ tells us how spread out a distribution is / how much the data points vary.

marks %
$D_1$ : 38, 40, 60, 65, 70, 70, 72, 75, 88, 90
Shannons class
dangerous Sadness

marks %
$D_2$ 65, 70, 70, 71, 75, 80, 80, 81
Fan class
↳ safe more consistant

* ☆ lower the measure of dispersion, the better

## Variance ($s^2$)

data point

mean

$$s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1} = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n-1}$$

total # of data points

P 1.14
1.15

| $x_i$ | $(x_i - 77)$ <br> $(x_i - \bar{x})$ | $(x_i - \bar{x})^2$ |
|---|---|---|
| 52 | $52 - 77 = -25$ | $(-25)^2 = 625$ |
| 59 | $59 - 77 = -18$ | $(-18)^2 = 324$ |
| 60 | ⋮ | ⋮ |
| ⋮ | | |
| 96 | $96 - 77 = 19$ | $(19)^2 = 361$ |
| | SNM | 3342 |

$$s^2 = \frac{SNM}{n-1} = \frac{3342}{18}$$

class A → $s^2 = 185.67$

class B → $s^2 = 304.5$

$s^2$

## Standard Deviation ($s$) ($s_x$)

(the square root of variance)

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$$

$$s = \sqrt{\frac{sum}{n-1}}$$

ex. pg 1.14

Di

$\bar{x} = 77\%$

$s^2 = 185.6$

$s = \sqrt{185.6}$

$s = 13.6\%$

The majority of the data points fall between $\bar{x} - s$ and $\bar{x} + s$.

$77 - 13.6 = 63.4\%$

$77 + 13.6 = 90.6\%$

} majority of marks ranging here

P 1.20 Instructions for entering a distribution

$D_1$ / Class A

P 1.14

$\bar{x} = 77\%$

$s^2 = 185.67$

$s = 13.60\%$

enter: $D_2$ / Class B

P) 1.14

$\bar{x} = 77\%$

$s^2 = 304.5$

$s = 17.4\%$

P 1.27 - 1.31

Q: Which teacher is performing better?

teacher A because lower s (the better)

# Mean Deviation : (MD)

Recall

$|2| = 2$

absolute value of

$|-21| = 2$

$|-\infty| = \infty$

$$MD = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n} = \frac{|x_1 - \bar{x}| + |x_2 - \bar{x}| + \cdots + |x_n - \bar{x}|}{n}$$

$\bar{x} = 70$

| $x_i$ | $x_i - \bar{x}$ | $\lvert x_i - \bar{x} \rvert$ |
|---|---|---|
| 70 | $70 - 70 = 0$ | $\lvert 0 \rvert = 0$ |
| 72 | $72 - 70 = 2$ | $\lvert 2 \rvert = 2$ |
| 70 | | |
| ⋮ | | ⋮ |
| 69 | $69 - 70 = -1$ | $\lvert -1 \rvert = 1$ |
| | Sum | 8 |

$MD = \frac{\text{sum}}{n}$

$MD = \frac{8}{7} = 1.14$

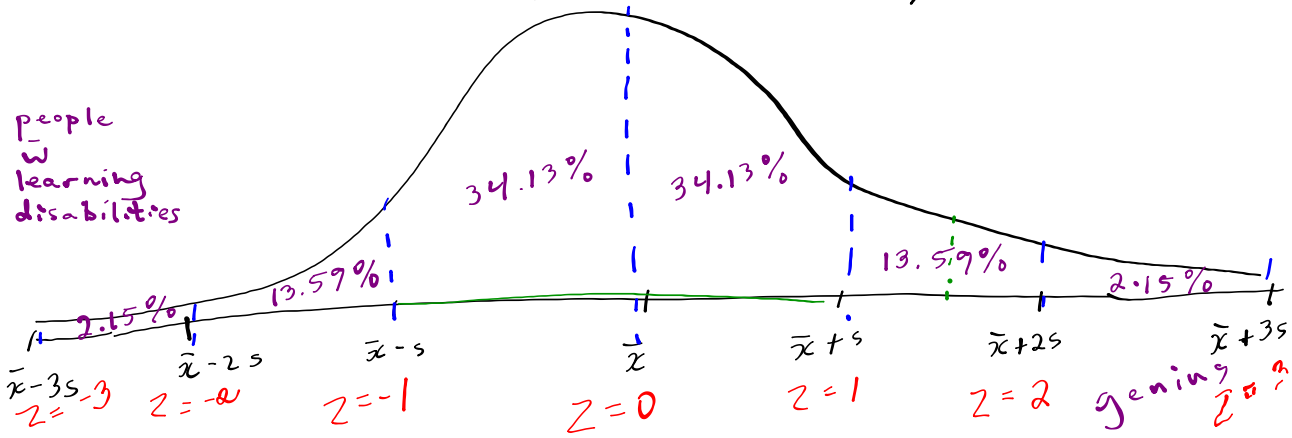
? 1.32

the majority of the marks range from $70 - 1.14 = 68.86$

$70 + 1.14 = 71.14$

# Unit 2: Standard Score (z)

- a # inbetween -3 and 3
- tells you how many standard deviations a data point is from the mean ($\bar{x}$)



1

$$Z = \frac{x_i - \bar{x}}{s}$$

✡ Nota Bene : each data point has a Z-score

P 2.5

Class 536 :

56, 56, 57, 58 ... 84, 92

Simon got 84% . find Simon's Z-score

$\bar{x} = 68.4\%$

$s = 10.46\%$

$$Z = \frac{x_i - \bar{x}}{s}$$

Shannon's in the same class

she got 92%

What's her Z-score

$Z = 2.25$

$$Z = \frac{84 - 68.4}{10.46}$$

$Z_s = 1.49$

Who performed better?

Joel and Josh are in different classes

Joel: 70%

$Z = 2$

Avg = ?

Josh = 75%

$Z = 0$

Avg = 75%

hard class

1 - unknown

1 - equation

$Z = \frac{x_i - \bar{x}}{s}$

- Who performed better? → Joel
- How could you explain Josh's higher mark but lower Z-score?

## Typical Exam Question

$x_i = 90\%$

$s = 10\%$

$Z = 2.7$

What's the class average? $\bar{x}$

$$Z = \frac{x_i - \bar{x}}{s}$$

$$10(2.7) = \left(\frac{90 - \bar{x}}{10}\right)10$$

$$27 = 90 - \bar{x}$$

$$-63 = -\bar{x}$$

$$\bar{x} = 63\%$$

BEDMAS

solve for $\bar{x}$ by doing opposite operations to both sides

John's $Z = 1.7$

Class average = 72%

Standard deviation = 5%

What's John's mark?

Typical Exam Question

**Student A**

$z_A = 2$

$x_i = 80$

$\bar{x} = 60$

$z = \frac{x_i - \bar{x}}{s}$

$2 = \frac{80 - 60}{s}$

$2 = \frac{20}{s}$

$2s = 20$

$s = 10$

**Student B**

$z_B = 1$

$x_i = 80$

$\bar{x} = 60$

$z = \frac{x_i - \bar{x}}{s}$

$1 = \frac{80 - 60}{s}$

$1 = \frac{20}{s}$

$s = 20$

∴ Class B has a higher standard deviation

Student A  Student B

$Z_A > Z_B$

$\frac{x_A - \bar{x}_A}{s_A} > \frac{x_B - \bar{x}_B}{s_B}$

$\frac{x_i - \bar{x}}{s_A} > \frac{x_i - \bar{x}}{s_B}$

$\frac{s_B(x_i - \bar{x})}{(x_i - \bar{x})} > \frac{s_A(x_i - \bar{x})}{(x_i - \bar{x})}$

$s_B > s_A$

* inequality sign flips when dividing or multiplying by a negative #

Hmwk:

p. 1.27 - 1.31

p. 1.39 #1 and #2

p 2.8 - 2.9

p 2.11 - 2.13