Math 200 Notes 12 4 2023. Please take detailed notes.

Section 9.5: Measures of Dispersion:

S_1 : 4, 6, 8, 10, 12, 14, 16. (We're getting from some experiment.)

S_2 : 6, 7, 9, 10, 11, 13, 14 (Another run of the experiment, and we^2 have some more data values.)

average=mean = $\frac{4+6+8+10+12+14+16}{7} = \frac{70}{7} = 10$ (measure of central tendency)

range=spread of the data= max-min= $16-4=12$ (extent to which data is spread out)

average =mean for second data set= $\frac{6+7+9+10+11+13+14}{7} = \frac{70}{7} = 10$

range= $14-6=8$ (less spread out than the first one)

Summary: both runs of the experiment(particular of the experiment don't matter) have the same average value of 10, but the second run has less spread out.

averages =means= $\frac{\text{add up all your values}}{\text{number of values}}$ range= max value-min value

Averages are affected by outliers: Experiment gives us 1, 2, 1, 3, 4, 5, 25 (25 is outlier)

Big idea: When outliers are involved , averages are not good measures of central tendency b/c they are affected by the outliers, and sometimes this is drastic. In this example, 5.9 is greater than 6 of the 7 values.

b/c averages have sums in tops: $\frac{1+2+1+3+4+5+25}{7} = \frac{41}{7} = 5.9$

5.9>1, 5.9>2, 5.9>1, 5.9>3, 5.9>4,5.9>5, 5.9<25

Median: middle value after we arrange our data set from low to high or high to low!

example 3 in book: $\underbrace{2, 2, 3, 4}_{4}, 5, \underbrace{7, 7, 7, 11}_{4}$ (already arranged from low to high)

Notice on either side of the 5 we have 4 values.

median=5 (in data set)

example: $\underbrace{2, 2, 3, 3}_{4}, \underbrace{4, 5}, \underbrace{7, 7, 7, 11}_{4}$ (10 values, so no middle value can be picked)

median= average of two middle values= $\frac{4+5}{2} = \frac{9}{2} = 4.5$

(not in data set 4.5)

Medians are NOT affected by outliers: 1, 2, 3, 1, 2, 3, 4, 5, 50 (50 is clearly different from the other values)

arrange the data: 1, 1, 2, 2, 3, 3, 4, 5, 50

median is 3. (50 doesn't affect this value)

instead: 1, 1, 2, 2, 3, 3, 4, 5, 1000

Mode: value that occurs with the greatest frequency

example: we do an experiment and we gather this data: 2, 3, 4, 5, 7, 15. In this case, there IS NO mode b/c each value occurrs only once. We do another experiment and get this data: 2, 2, 2, 3, 3, 7, 7, 7, 11, 15

2 is a mode. 7 is a mode.

This is a bimodal data set.

summary: mean=average= $\frac{\text{sum of values}}{number \text{ of values}}$, affected by outliers

median= arrange data and pick out middle value, *not* affected by outliers

range= max value-min value (spread of data)

mode= most frequently occuring value

percent change: $\frac{\text{final value-initial value}}{\text{initial value}} \cdot 100$

example: Imagine the average value of a data set goes from 10 to 11. Find the percent change in the average.

$\frac{11-10}{10} = \frac{1}{10} = 0.1 = 10\%$

example: Imagine median home price on a block in 2009 was 450,000, and in 2010 it rose to 475,000.

Find the percent change in the median home price: $\frac{475000-450000}{450000} = \frac{25000}{450000} = \frac{25}{450} = 0.056$ or 5.6% increase.

symbols for averages: $\bar{x} \Leftarrow$ read as "x bar" example: what's the average of 1, 1, 1, 1, 1, 1 (from some experiment)

$\bar{x} = \frac{1+1+1+1+1+1}{6} = \frac{6}{6} = \mathbf{1}$ (average is 1 and its equal to each value in the list)

range= $1-1=0$ (no spread in the data)

median= average out two middle values= $\frac{1+1}{2} = \frac{2}{2} = \mathbf{1}$

mode= 1

example: Imagine we burn 100 calories one day. We burn 200 calories the next day. We burn 300 calories the last day.

aveage= $\frac{100+200+300}{3} = \frac{600}{3} = 200$ calories per day! A way to think about this: Shift 100 calories from day 3 to day 1 and then you have 200, 200 $and \cdot$ 200 and still we get 200+200+200=600 calories in total.

deviations: Imagine we do experiment and get this data set: 1, 2, **3**, 1, 4, 5, 6, 2

average= $\bar{x} = \frac{1+2+3+1+4+5+6+2}{8} = \frac{24}{8} = 3$

deviations from average:

$1-3=-2 \Leftarrow$ 1 is 2 units below the average of 3

$2-3=-1 \Leftarrow$ 2 is 1 unit below the average of 3

$\mathbf{3}-\mathbf{3}=0 \Leftarrow \mathbf{3}$ is 0 units below or above the average.

1-3= $-2 \Leftarrow$ 1 is 2 units below the average of 3

$4-3=+1 \Leftarrow$ 1 unit above the average of 3

5-3= $+2 \Leftarrow$ 2 units above the average of 3

$6-3=3 \Leftarrow$ 3 units above the average of 3

$2-3=-1 \Leftarrow$ 1 units below the average of 3

observation: add up the deviations: $-2-1+0-2+1+2+3-1=\mathbf{0}$ Perhaps this is true just for this data set.
For **ANY** data set, the **sum** of deviations is 0.

example: Professor Bob tested 8 students in the class . The scores were 100,94,85, 79, 70, 69, 65 and 62.

average score: $\frac{100+94+85+79+70+69+65+62}{8} = \frac{624}{8} = 78 \Leftarrow$ average test score!

deviations from average: $100-78=22$ (22 points above average score)

$94-78=16$ points above average score

$85-78=7$ points above average score

$79-78=1$ point above average score

$70-78=-8$ (8 points BELOW the average)

$69-78=-9$ (9 points below the average)

$65-78=-13$ (13 points below the average)

$62-78=-16$ (16 pts below the average)

sum of deviations:
$22+16+7+1-8-9-13-16$

$= \mathbf{0}$ (again equal to 0)

When we use letters in math, x or y, or z , μ, ν, α, β , Ω

How many ways are there to write the number 1?? $2-1$, $\frac{4}{4}$, $\frac{5-4}{5-4}$, and so on, so there are ∞ 'ly many ways to write any value, and not just 1, in math. Symbols in math have no inherent meaning. It's only the meaning we give them they have.