Data Analysis
Part I

Measures of Central Tendency

Mean (Average)

- Population Mean: $\mu = \dfrac{\sum_{i=1}^{N} x_i}{N}$

- Sample Mean: $\bar{X} = \dfrac{\sum_{i=1}^{n} x_i}{n}$

Median

- The data must be ranked (sorted in ascending order) first. The median is the number in the middle.

Mode

- The mode is the most frequent data value. There may be no mode if no one value appears more than any other.

The eight observations are $x_1 = 12.6$, $x_2 = 12.9$, $x_3 = 13.4$, $x_4 = 12.3$, $x_5 = 13.6$, $x_6 = 13.5$, $x_7 = 12.6$, and $x_8 = 13.1$. The sample mean is

$$\overline{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{\sum_{i=1}^{8} x_i}{8} = \frac{12.6 + 12.9 + \cdots + 13.1}{8} = \frac{104}{8} = 13.0$$

Median & Mode

Data: 12, 5, 9, 8, 4

Arrange data in ascending order: 4, 5, 8, 9, 12

- Median: 8
- Mode: No mode. Don't write mode as 0

Measures of Dispersion

Range

- The range is the simplest measure of variation to find. It is simply the highest value minus the lowest value.

RANGE = MAXIMUM - MINIMUM

Data: 12, 5, 9, 8, 4

Range: 12-4 = 8

Variance

- "Average squared deviation from the mean"

$$Population\ Variance\ =\ \sigma^2\ =\ \frac{\sum (x - \mu)^2}{N}$$

$$Sample\ Variance\ =\ s^2\ =\ \frac{\sum (x - \bar{x})^2}{n - 1}$$

Standard Deviation

Population Standard Deviation,

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}}$$

Sample Standard Deviation

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

The eight observations are $x_1 = 12.6$, $x_2 = 12.9$, $x_3 = 13.4$, $x_4 = 12.3$, $x_5 = 13.6$, $x_6 = 13.5$, $x_7 = 12.6$, and $x_8 = 13.1$.

i	x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
1	12.6	−0.4	0.16
2	12.9	−0.1	0.01
3	13.4	0.4	0.16
4	12.3	−0.7	0.49
5	13.6	0.6	0.36
6	13.5	0.5	0.25
7	12.6	−0.4	0.16
8	13.1	0.1	0.01
	104.0	0.0	1.60

$$s^2 = \frac{1.60}{8 - 1} = \frac{1.60}{7} = 0.2286$$

$$s = \sqrt{0.2286} = 0.48$$

Grouped data: Mean and Standard deviation

- Mean from Frequency Distribution:

$$\bar{X} = \frac{\sum mf}{\sum f}$$

Standard deviation, s

$$= \sqrt{\frac{\sum(fm^2) - \frac{(\sum fm)^2}{n}}{n-1}}$$

Data:
5,6,7,9,10.5,11.5,10,11,9.5,13,14,15,16,16.5,17,17.5,18,18.5,19,19.5

Class Interval	Freq (f)	Class Mid point (m)	fm	fm^2
4 -8	3	6	18	108
8-12	6	10	60	600
12-16	4	14	56	784
16-20	7	18	126	2268
	n=20		260	3760

$$\bar{X} = \frac{\sum mf}{\sum f} = \frac{260}{20} = 13$$

$$s = \sqrt{\frac{\sum(fm^2) - \frac{(\sum fm)^2}{n}}{n-1}} = \sqrt{\frac{3760 - \frac{(260)^2}{20}}{20-1}} = 20$$

Exercise 1

Compute mean, median, mode and standard deviation of

I] 8, 9, 12, 5, 7

II]

Class Interval	Freq (f)	Class Mid point (m)	fm	fm^2
5 -11	8			
11-17	12			
17-23	14			
23-29	5			
	n=			

Solution Exercise 1

No. I	
Mean	8.2
Std Dev	2.588436
Mode	No
Median	8

No. II					
Class Interval		Freq (f)	Class Mid point (m)	fm	fm^2
5	11	8	8	64	512
11	17	12	14	168	2352
17	23	14	20	280	5600
23	29	5	26	130	3380
		n=39		642	11844

Mean	16.46
Std dev	5.79
Mode	Interval 17-23
Median	Interval 11-17