Chapter 3

3.1 Measures of Location:

(1) Mean (2) Median (3) Mode (4) Percentiles (5) Quartile

(2) Median: وسيط — (4) Percentiles: مئين — (5) Quartile: ارباع

(1) Mean: The most important measure of location is the mean or average value for a variable.

Sample Mean: $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ is a sample statistics.

n: is the number of observations

x_i: is the i^{th} observation

Example one: Consider the following sample of a class marks:

46	54	42	46	32
x_1	x_2	x_3	x_4	x_5

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{46+54+42+46+32}{5} = 44$$

The sample mean class mark is 44.

Population Mean $\mu = \frac{\sum_{i=1}^{N} x_i}{N}$ N is the total (whole) population number

Note that the sample mean $\bar{x}$ is a point estimator of the population mean μ.

[2] **Median:** is a measure of central location.

→ Arrange the data in ascending order (smaller to largest value)

(a) For an odd number of observations, the median is the middle value.

(b) For an even = = = , = = = = average of the two middle values.

In Example one: Arranging the data ascending order:

32 42 46 46 54

Because n = 5 odd, the median is the middle value = 46

Example two: If the data were ordered like:

32 42 46 46 54 60 70 80 90 96

Then n = 10 even, the median = $\frac{54+60}{2} = \frac{114}{2} = 57$

---

* Whenever a data set contains extreme values, the median is often the prefered measure of central location

---

[3] **Mode:** is another measure of location.

The mode is the value that occurs with greatest frequ

In Example one: The mode is 46

---

* If data contains exactly two modes, then the data are called bimodal.

* If = , more then = = , = = = = = multimodal.

* More than 2 modes would not be helpful in measuring location for data.

[4] Percentile (المئين) :

The p^{th} percentile is a value such that at least p percent of the observations are less than or equal to this value, and at least (100 - p) percent of the observation are greater than or equal to this value.

How to calculate the p^{th} percentile ?

step 1 : Arrange the data in ascending order (smallest to greatest value).

step 2 : Compute an index $i = \left(\frac{p}{100}\right) n$ where

p is the percentile of interest

n is the number of observations.

step 3 : [a] if i is not integer, then round up to the next integre greater than i . This tells us the position of the p^{th} percentile.

[b] if i is an integer , then the p^{th} percentile is the average of the values in positions i and i+1.

Example [a] Compute the 85^{th} percentile for the following data

46 54 42 46 32

step 1 : 32 42 46 46 54

step 2 : $i = \frac{85}{100} \times 5 = 4.25$

step 3 : Because i is not integer $\Rightarrow i = 5$ 5^{th} position $\Rightarrow$ we see that the 85^{th} percentile is the data value 54

[b] Compute the 40^{th} percentile

$i = \frac{40}{100} \times 5 = 2$ integer $\Rightarrow$ the 40^{th} percentile $= \frac{42+46}{2} = 44$

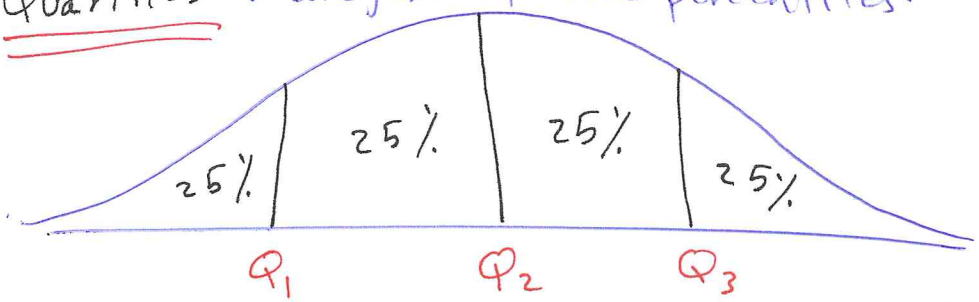
(c) Compute the 50^{th} percentile:

$$i = \left(\frac{50}{100}\right) \times 5 = 2.5 \Rightarrow i = 3$$

$\Rightarrow$ the 50^{th} percentile is 46 "Median"

Note that the 50^{th} percentile is also the median.

(5) Quartiles: are just specific percentiles.



Q_1: first Quartile (25^{th} percentile)

Q_2: Second Quartile (50^{th} percentile) (Median)

Q_3: Third Quartile (75^{th} percentile)

Example: Compute Q_1, Q_2, Q_3 for the data:

46 54 42 46 32

Step 1: 32 42 46 46 54

Step 2:

- $Q_1 = 25^{th}$ percentile

$$i = \left(\frac{25}{100}\right) \times 5 = 1.25 \Rightarrow i = 2 \text{ } 2^{nd} \text{ position}$$

$\Rightarrow Q = 42$

- $Q_2 = 50^{th}$ percentile = median = 46 see part (c)

- $Q_3 = 75^{th}$ percentile

$$i = \left(\frac{75}{100}\right) \times 5 = 3.75 \Rightarrow i = 4 \text{ } 4^{th} \text{ position}$$

$\Rightarrow Q_3 = 46$