UNIT 17 *Measures of Central Tendency* Introduction to SIM

Learning objectives

After completing this unit of STUDENT INSTRUCTIONAL MATERIAL (SIM) you should be able to

- calculate the mean, mode, median and range for a set of discrete data
- use tally charts and tables to calculate the mean of a data set
- undertake calculations with the mean
- calculate or estimate mean, mode and median from a set of grouped data.

All sections in this unit are on the CXC/CSEC mathematics syllabus; they also provide an important foundation for further work in Statistics, both in the next unit and in further study.

Key points and principles

- 'Mean', 'median' and 'mode' are all measures of the *average* of a set of data; we call them
  measures of central tendency
- You should realise that, although the mean, median and mode each represent an average for the data set, it is important to understand the context in any particular case, in order to appreciate which one is the most appropriate to use.
- The 'range' is one measure of the spread or variation of the data in a data set; we call this a *measure of variation*.

Facts to remember

- To find the *median*, you need first to put the data into numerical order and then to identify the middle value (if there is an odd number of values) or take the mean of the middle two values (if there is an even number of values).
- The *mode* is the value that has the highest frequency.

Glossary of terms

Mean - the arithmetic mean, obtained by adding together all the numbers in a set of data and dividing by how many numbers there are in the set.

For example, the mean of a set of numbers, $x_1, x_2, \ldots, x_n$ is calculated from the formula

$$\text{mean} = \frac{x_1, x_2, \ldots, x_n}{n}$$

So, for the data set $\{ 5, 7, 2, 5, 1, 2, 3, 5, 6 \}$,

$$\text{the mean value} = \frac{(5 + 7 + 2 + 5 + 1 + 2 + 3 + 5 + 6)}{9} = \frac{36}{9} = 4.$$

UNIT 17 *Measures of Central Tendency* Introduction to SIM

Median - the median is the middle value when the values are listed in numerical order.

For example, for the data set { 5, 7, 2, 5, 1, 2, 3, 5, 6 }, the numerical order is

$$1,\ 2,\ 2,\ 3,\ 5,\ 5,\ 5,\ 6,\ 7$$

↑
middle value

and so the median $= 5$.

Note that if there are an even number of values in the data set, the median is the mean of the middle two values.

For example, for the data set { 5, 7, 2, 5, 1, 2, 3, 5 }, the numerical order is

$$1,\ 2,\ 2,\ \underbrace{3,\ 5,}\ 5,\ 5,\ 7$$

↑
middle value

and so the median $= \dfrac{(3+5)}{2} = 4$.

Mode - the mode is the numerical value that has the highest frequency in a set of data, that is, the value that occurs most often.

For example, for the data set { 5, 7, 2, 5, 1, 2, 3, 5, 6 }, you can write the values as

Value	1	2	3	4	5	6	7
Frequency	1	2	1	0	3	1	1

The mode is 5, as it has the highest frequency (3).

Range - the range is the difference between the highest and the lowest values in the data set.

For example, for the data set { 5, 7, 2, 5, 1, 2, 3, 5, 6 },

the range is $7 - 1 = 6$.