CH. 2 DESCRIBING DATA WITH NUMERICAL MEASURES

Def: PARAMETERS ARE MEASURES ASSOCIATED WITH POPULATION

STATISTICS " " " SAMPLE

§2.2 MEASURES OF CENTER

3 WAYS

1. ARITHMETIC MEAN (AVERAGE)

DENOTED $\bar{x}$ FOR SAMPLE, μ FOR POPULATION

FOR SAMPLE WITH n MEASUREMENTS, $x_1, x_2, \dots, x_n$,

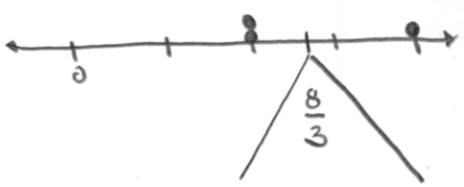
$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

$$\boxed{\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}} = \frac{\sum x_i}{n}$$

$\downarrow$

WE CAN USE THIS TO ESTIMATE μ.

$\frac{2+2+4}{3} = \frac{8}{3}$ "BALANCE POINT"



2. MEDIAN

Def: MEDIAN m OF A SET OF n MEASUREMENTS IS THE VALUE x THAT FALLS IN THE MIDDLE POSITION WHEN THE MEASUREMENTS ARE ORDERED FROM SMALLEST TO LARGEST.

e.g. 1 1 3 **4** 8 12 20 (n ODD)

$m = 4$

e.g. 1 2 5 **8 12** 12 14 16 (n EVEN)

$$m = \frac{8+12}{2} = 10$$

Note: ROUGHLY HALF MEASUREMENTS ABOVE/BELOW MEDIAN.

Note: POSITION OF MEDIAN IS $\frac{n+1}{2}$

(IF NOT INTEGER, TAKE AVERAGE OF TWO ADJ. VALUES)

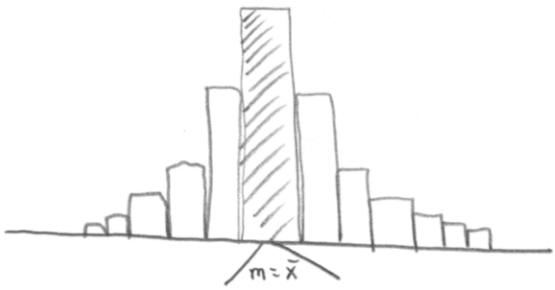
REMARK: MEDIAN IS LESS SENSITIVE TO OUTLIERS (i.e. EXTREME VALUES)

4 4 4 5 **5 5** 5 6 6 6 $\quad m = 5 \quad \bar{x} = 5$

4 4 4 5 **5 5** 5 6 6 16 $\quad m = 5 \quad \bar{x} = 6$

(MEDIAN INCOME)

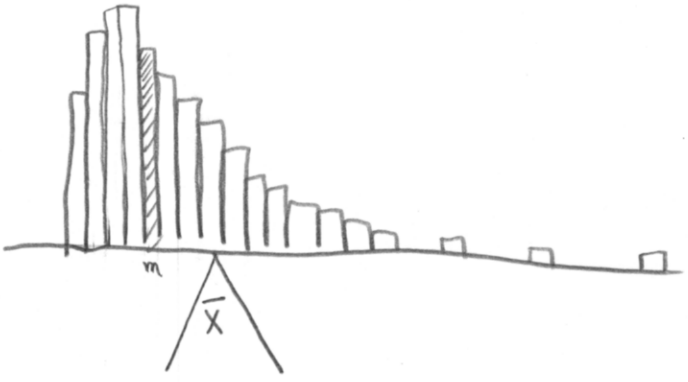
SYMMETRIC



MEAN = MEDIAN

BALANCE POINT = MIDDLE VALUE

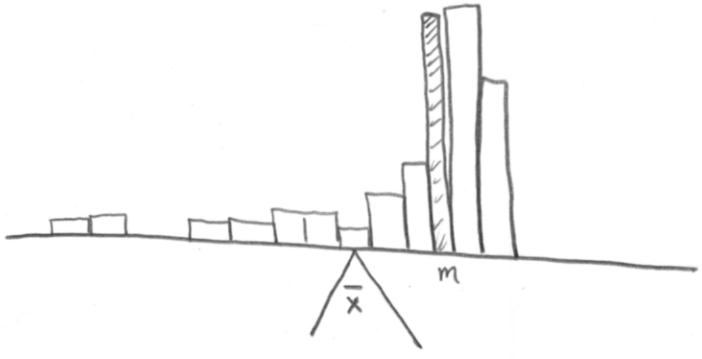
SKEWED RIGHT



MEAN > MEDIAN

$\bar{x} > m$

SKEWED LEFT

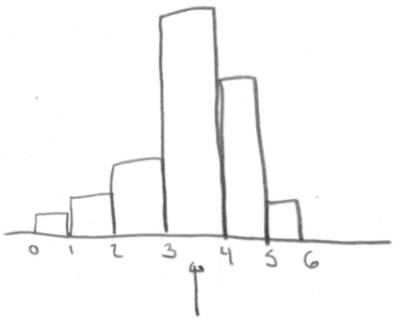


$\bar{x} < m$

3. MODE

Def: THE MODE IS THE CATEGORY THAT OCCURS MOST FREQUENTLY, OR THE MOST FREQUENTLY OCCURRING VALUE OF x.

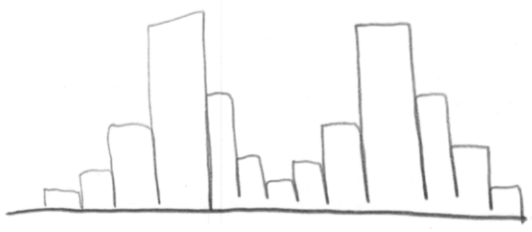
WHEN MEASUREMENTS OF A CONTINUOUS VARIABLE HAVE BEEN GROUPED AS A FREQUENCY OR RELATIVE FREQUENCY HISTOGRAM, THE CLASS WITH THE HIGHEST PEAK IS CALLED THE MODAL CLASS, AND THE MIDPOINT OF THAT CLASS IS TAKEN TO BE THE MODE.



MODAL CLASS: $3 \leq x < 4$

MODE = 3.5

Note: THERE CAN BE MORE THAN ONE MODE



MALE / FEMALE

YOUNG / OLD