Statistics

Homework 1 – Representation of Data

Abraham Murciano

March 9, 2020

1. (a) For the following data set,

   $$1, 8, 1, 5, 8, 6, 3, 3, 3, 7$$

   the mid-range, average, median and mode are as follows.

   $$\text{Mid-range} = \frac{1+8}{2} = 4.5$$
   $$\text{Average} = \frac{1+8+1+5+8+6+3+3+3+7}{10} = 4.5$$
   $$\text{Median} = \frac{3+5}{2} = 4$$
   $$\text{Mode} = 3$$

   (b) For the following data set,

   $$14, 18, 30, 31, 15, 18, 27$$

   the mid-range, average, median and mode are as follows.

   $$\text{Mid-range} = \frac{14+31}{2} = 22.5$$
   $$\text{Average} = \frac{14+18+30+31+15+18+27}{7} \approx 21.86$$
   $$\text{Median} = 18$$
   $$\text{Mode} = 18$$

3. In the month of August 2010 it was very hot. The Israeli meteorological service recorded the maximum temperature. The results are presented in Table 1.

Number of days	Max. temperature (°C)
2	27-29
6	30-32
12	33-35
5	36-38
4	39-41

Table 1: Maximum daily temperature frequencies for August 2010

Max. temp. (°C)	Frequency	Relative freq.	Cumulative freq.	relative cumulative freq.
27-29	2	0.0690	2	0.0690
30-32	6	0.2069	8	0.2759
33-35	12	0.4138	20	0.6897
36-38	5	0.1724	25	0.8621
39-41	4	0.1379	29	1.0000

Table 2: Many types of frequencies of the data in Table 1

(a) Table 2 shows the frequency, relative frequency, cumulative frequency, and relative cumulative frequency.

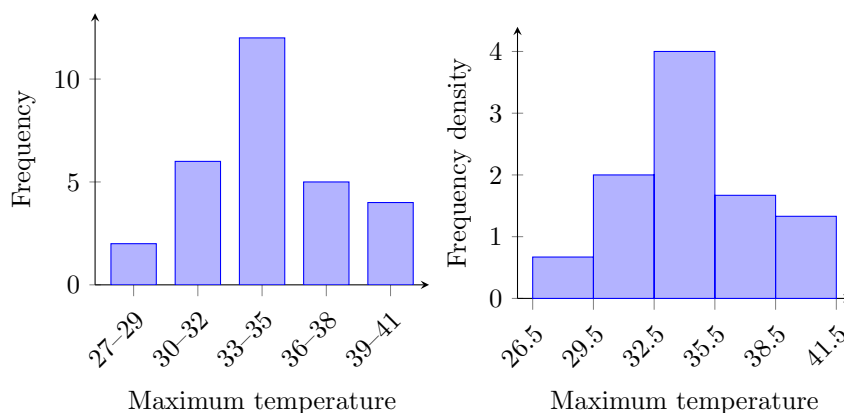(b) Figure 1 shows a bar graph of the data in Table 1 as well as a histogram.

Figure 1: Bar graph of the data in Table 1 (left) and a histogram of the same (right)

(c) The mid-range, average, median and mode for this data are as follows.

$$\text{Mid-range} = \frac{28 + 40}{2} = 34$$
$$\text{Average} = \frac{2 \times 28 + 6 \times 31 + 12 \times 34 + 5 \times 37 + 4 \times 40}{29} \approx 34.3103$$
$$\text{Median} = \frac{3\left(\frac{29}{2} - 8\right)}{12} + 32.5 = 34.125$$
$$\text{Modal class} = [32.5, 35.5)$$

6. (a) For two classes with the same number of students, the average math grade of all the students combined is equal to the average of the average grades of each class separately. As an example, consider the data for the two classes in Table 3.

The average grade of Class 1 is

$$\frac{100 + 98 + 86 + 79}{4} = 90.75$$

The average grade of Class 2 is

$$\frac{86 + 68 + 66 + 59}{4} = 69.75$$

| | Class 1 |
Student	Grade
A	100
B	98
C	86
D	79

| | Class 2 |
Student	Grade
E	86
F	68
G	66
H	59

Table 3: The grades of two classes of equal size

The average of these two averages is

$$\frac{90.75 + 69.75}{2} = 80.25$$

This is the same as the average grade of all eight students.

$$\frac{100 + 98 + 86 + 86 + 79 + 68 + 66 + 59}{8} = \frac{642}{8} = 80.25$$

(b) The median math grade for two classes of equal size when combined into one big class is **not** equal to the average of the two medians. For example, consider the data for two classes in Table 3.

The median of Class 1 is 92, the median of Class 2 is 67, and the average of the two medians is 79.5. However, the median of all eight students is 82.5.

(c) The modal math grade of the two classes when combined into one big class is **not** necessarily less than or equal to the modal grade of each one. As an example, if two classes have grades as shown in Table 4, Class 1 has a mode of 100, but that is greater than the mode of the two classes combined, which is 60.

| | Class 1 |
Student	Grade
A	100
B	100
C	85
D	82

| | Class 2 |
Student	Grade
E	74
F	60
G	60
H	60

Table 4: The grades of two classes of equal size

(d) It is true that the mid-range math grade for the two classes when combined into one larger class must always lie between the mid-ranges of each of the two classes.

As an example, consider the data in Table 4. The mid-range for Class 1 is

$$82 + \frac{100 - 82}{2} = 91$$

The mid-range for Class 2 is

$$60 + \frac{74 - 60}{2} = 67$$

And the mid-range for the two classes combined is

$$60 + \frac{100 - 60}{2} = 80$$

which is between 67 and 91.

7. Consider the data in Table 5.

x	$f(x)$
40-50	18
50-60	6
60-70	k
70-80	9
80-90	10
90-100	5

Table 5: Partial data for a variable x

(a) It is possible to calculate the median value for x, since we know for certain that the median class is the 60-70 class, and there are exactly 24 observations before and 24 observations after that class. Therefore the median value is precisely the midpoint of the median class. That is 65.

(b) If the average is equal to 65.4, then the missing value in Table 5 must satisfy this equation.

$$\frac{18 \times 45 + 6 \times 55 + 9 \times 75 + 10 \times 85 + 5 \times 95 + 65 \times k}{18 + 6 + 9 + 10 + 5 + k} = 65.4$$

Now we can solve for k to obtain

$$\frac{3140 + 65n}{48 + k} = 65.4$$
$$3140 + 65n = 65.4(48 + k)$$
$$3140 + 65n = 3139.2 + 65.4k$$
$$3140 - 3139.2 = 65.4k - 65n$$
$$0.8 = 0.4k$$
$$k = 2$$

(c) If the 90th percentile is 89.4, the missing value in the table must satisfy this equation.

$$89.4 = \frac{(48 + k)\left(\frac{90}{100}\right) - (33 + k)}{10} \times (90 - 80) + 80$$

Then we are able to solve for k to obtain the missing value in the table.

$$9.4 = \frac{9}{10}(48 + k) - 33 - k$$
$$42.4 = \frac{9}{10}(48 + k) - k$$
$$42.4 = 43.2 + 0.9k - k$$
$$-0.8 = -0.1k$$
$$k = 8$$