9.1 Frequency Distributions and Measures of Central Tendency

Frequency Distributions

1. A survey asked a random sample of 30 business executives for their recommendations as to the number of college credits in management that a business major should have. The results are shown below. Group the data into intervals and find the frequency of each interval. Then create a frequency histogram and frequency polygon to visualize the data.

3	25	22	16	0	9	14	8	34	21
15	12	9	3	8	15	20	12	28	19
17	16	23	19	12	14	29	13	24	18

Credits	Tally	Frequency
0-4		3
5-9		4
10-14		6
15-19		8
20-24		5
25-29		3
30-34		1

Credits	Tally	Frequency
0-6		3
7-13		8
14-20		11
21-27		5
28-34		3

Note 1. You have artistic freedon in choosing the classes/intervals into which your data is partitioned. The only rule is that the class widths must all be the same. However, the most effective histograms have roughly 5-10 classes/intervals.

Measures of Central Tendency

Mean

The **mean** of the n numbers $x_1, x_2, x_3, \ldots, x_n$ is

$$\bar{x} = \frac{\Sigma x}{n}.$$

2. Find the mean of the data from question 1 in three ways:
   (a) using the raw data,
   (b) using the first frequency table, and
   (c) using the second frequency table.

Mean of a Grouped Distribution

The mean of a distribution, where x represents the midpoints, f the frequencies, and $n = \Sigma f$, is

$$\bar{x} = \frac{\Sigma xf}{n}.$$

Note 2. There are two different types of data, each with a different notations for the mean. **Population** data is gathered from every idividual of interest (e.g. the birth weight of every human ever), and **sample** data is gathered from a (hopefully) random sample taken from the population (e.g. the birthweight of 35 randomly selected humans). Regardless, the mean is calculated the same.

	Population	Sample
Mean	μ	$\bar{x}$

3. Two startup companies A and B each have 5 employees. The salaries of all employees are given below.

Company	Salaries				
A	\$24,000	\$52,000	\$76,000	\$83,000	\$144,000
B	\$24,000	\$52,000	\$76,000	\$83,000	\$144,000,000

   (a) Compute the means of salaries at companies A and B
   (b) Compute the medians of salaries at companies A and B

Definition 1. The **median** is defined as the middle entry in a set of data arranged in either *increasing or decreasing order*. If there is an even number of entries, the median is defined to be the mean of the two center entries.

More specifically, given n numbers arranged in order, the median is

   (a) the number in position $\frac{n+1}{2}$ if n is odd, and
   (b) the mean of the numbers in positions $\frac{n}{2}$ and $\left(\frac{n}{2}+1\right)$ if n is even.

Definition 2. The **mode** of a data set is the value(s) that appears most frequently. When data is summarized in a frequency table or histogram, the mode is defined as the midpoint(s) of the class(es)/interval(s) with the highest frequency.

4. Find the median and mode of the data in question 1.
   Hint: Don't forget to sort the data.