Statistics Review Solutions

1. Katrina must take five exams in a math class. If her scores on the first four exams are 71, 69, 85, and 83, what score does she need on the fifth exam for her overall mean to be at least 90?

   Solution:

   In order to determine the answer to this question, we would need to add all 5 exam scores together and divide the sum by 5. This needs to equal 90. Since we don't have all 5 exam scores, we need to go back a step and just figure out what all the exams need to add up to so that when we divide that number by 5, we get 90. In other words, we need to determine what number divided by 5 is equal 90. This number is 5 times 90 which is 450. All 5 exams need to add up to 450. The exams we have so far add up to 308 (71 + 69 + 85 + 83 = 300). We only have one exam left to get to 450, thus that exam score must be 142 (450 - 308 = 142). Thus it is not possible for Katrina to score high enough on the fifth exam to get an average of 90 (this is assuming that each test is worth 100 points).

2. Todd Booth, an avid jogger, kept detailed records of the number of miles he ran per week during the past year. The frequency distribution below summarized his records. Find the mean, median, and mode of the number of miles per week that Todd ran.

Miles Run per Week	Number of Weeks
0	5
1	4
2	10
3	9
4	10
5	7
6	3
7	4

   Solution:

   Although I am finding the mean and median without a calculator here, you will be able to use a calculator on the exam to find the mean, median, and standard deviation.

MEAN: To find the mean number of miles, you need to add all the numbers of miles from each week together and divide by the number of stores surveyed. The sum of all the numbers of miles is (Notice: since the number of weeks next to 0 is 5, I have to add in 0 five times.)
0 + 0 + 0 + 0 + 0 + 1 + 1 + 1 + 1 + 2 + 2 + 2 + 2 + 2 + 2 + 2 + 2 + 2 + 3 + 3 + 3 + 3 + 3 + 3 + 3 + 3 + 4 + 4 + 4 + 4 + 4 + 4 + 4 + 4 + 4 + 5 + 5 + 5 + 5 + 5 + 5 + 5 + 6 + 6 + 6 + 7 + 7 + 7 + 7 = 172
I now divide the sum by 52 since that is the number of weeks recorded. This will give me a mean of 172/52 = 3.307692308

MEDIAN: The median number of miles is the number of miles in the middle when all of the numbers are put in order. Since there are 52 numbers, the middle is in between the 26th and 27th number of miles. We will take these two numbers of miles, add them together and divide by two. The 26th number of miles is 3. The 27th number of miles is 3. Thus the median is (3 + 3)/2 = 3

MODE: The mode number of miles is the number of miles that occurs in the most weeks. This is easy to see from the table. The number of miles that occurs in the most weeks is 2 since there are 10 weeks with that number of miles and 4 since there are 10 weeks with that number of miles. Thus the data is bimodal with the number of miles 2 and 4 both being modes.

3. The mean salary of 12 men is $52,000, and the mean salary of 4 women is $84,000. Find the mean salary of all 16 people.
   Solution:
   To find a mean, we need to add all of the 16 salaries together and divide by 16. Since we don't have individual salaries, we need to settle for the totals amount made by the 12 men and the total amount made by the 4 women. We can find the total amount of the salaries of the 12 men by multiplying the mean salary of the 12 men by 12. This will give us a total of $624000. We can find the total amount of the salaries of the 4 women by multiplying the mean salary of the 4 women by 4. This will give us a total of $336000. If we add these two totals together, we get the sum of salaries of all 16 people. This sum is $960000. Now all we need to do to find the mean is divide this total

by 16. This will give us a mean of $60000 for all 16 people.

4. The frequency distribution below lists the results of a quiz given in Professor Gilbert's statistics class.

Score	Number of Students
10	5
9	10
8	6
7	8
6	3
5	2

a. Find the mean and standard deviation of the scores.
b. What percent of the data lies within one standard deviation of the mean?
c. What percent of the data lies within two standard deviations of the mean?
d. What percent of the data lies within three standard deviations of the mean?
e. Draw a histogram to illustrate the data.

Solution:

a. These we will not calculate by hand (although we could use either of the formulas given in the notes to do these calculation). According to the calculator, the mean of the scores is 8 and the standard deviation of the quiz scores is 1.435481125 which we will round to 1.4355.

b. To find what percent of the data falls within one standard deviation of the mean, we need to find how many of the students have quiz scores within one standard deviation of the mean. What this mean is that we need to find a quiz score one standard deviation below the mean (8 - 1.4355 = 6.5645) and a quiz score one standard deviation above the mean (8 + 1.4355 = 9.4355). We now want to put a mark on the table to show where 6.5645

happens (between 6 and 7) and a mark on the table to show where 9.4355 happens (between 9 and 10). Since I can't mark the table, I will shade in the number of stores that would fall between the two marks.

Score	Number of Students
10	5
9	10
8	6
7	8
6	3
5	2

We now need to add together the shaded numbers to find how many students have quiz scores which fall within one standard deviation of the mean. This number is 10 + 6 + 8 = 24. To find the percentage of scores that fall within one standard deviation of the mean, we will take 24, divide it by 34 (since there are 34 total student scores) and multiply that by 100 (to change from a decimal to a percent). This will give us (24/34)*100= 70.6%.

c. To find what percent of the data falls within two standard deviations of the mean, we need to find how many of the students have quiz scores within two standard deviations of the mean. What this means is that we need to find a quiz score two standard deviations below the mean (8 - 1.4355 - 1.4355 = 5.129) and a quiz score two standard deviations above the mean (8 + 1.4355 + 1.4355 = 10.871). We now want to put a mark on the table to show where 5.129 happens (between 5 and 6) and a mark on the table to show where 10.8713 happens (above 10). Since I can't mark the table, I will shade in the number of stores that would fall between the two marks.

Score	Number of Students
10	5
9	10
8	6
7	8
6	3
5	2

We now need to add together the shaded numbers to find how many students have quiz scores which fall within two standard deviations of the mean. This number is 5 + 10 + 6 + 8 + 3 = 32. To find the percentage of scores that fall within two standard deviations of the mean, we will take 32, divide it by 34 (since there are 34 total exam scores) and multiply that by 100 (to change from a decimal to a percent). This will give us (32/34)*100= 94.12%.

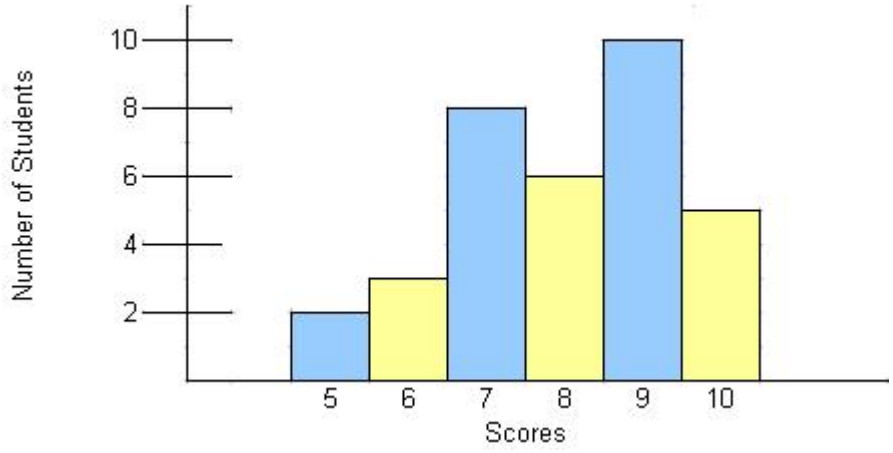
d. To find what percent of the data falls within three standard deviations of the mean, we need to find how many of the students have quiz scores within three standard deviations of the mean. What this means is that we need to find a quiz score three standard deviations below the mean (8 - 1.4355 - 1.4355 - 1.4355 = 3.6935) and a quiz score two standard deviations above the mean (8 + 1.4355 + 1.4355 + 1.4355 = 12.3065). We now want to put a mark on the table to show where 3.6935 happens (below 5) and a mark on the table to show where 12.3065 happens (above 10).

Since I can't mark the table, I will shade in the number of stores that would fall between the two marks.

Score	Number of Students
10	5
9	10
8	6
7	8
6	3
5	2

We now need to add together the shaded numbers to find how many students have quiz scores which fall within two standard deviations of the mean. This number is 5 + 10 + 6 + 8 + 3 + 2 = 34. To find the percentage of scores that fall within two standard deviations of the mean, we will take 34, divide it by 34 (since there are 34 total exam scores) and multiply that by 100 (to change from a decimal to a percent). This will give us (34/34)*100= 100%.

e.

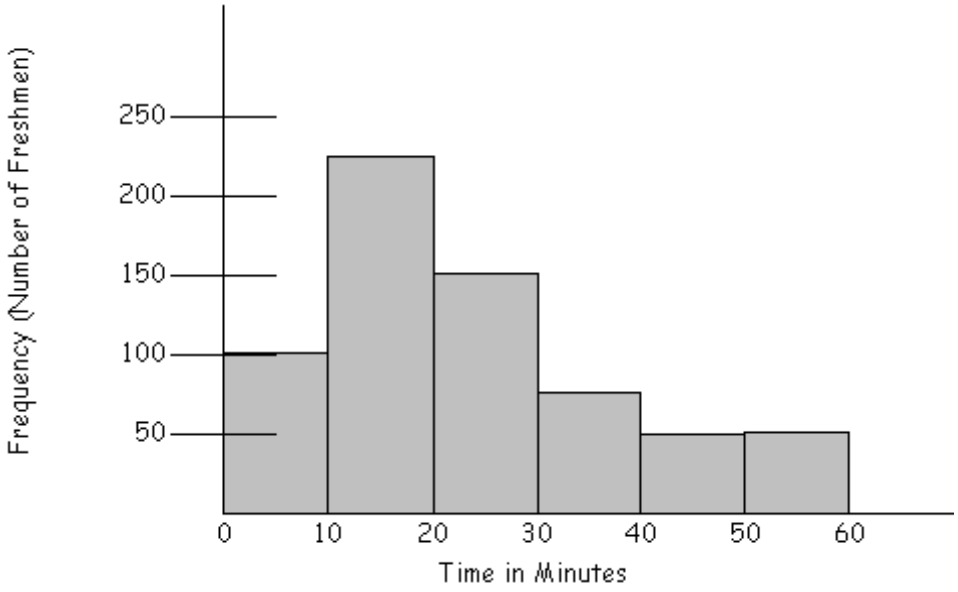


5. To examine the effects of a new registration system, a campus newspaper asked freshmen how long they had to wait in a registration line. The frequency distribution is given below. Complete the frequency distribution and draw a histogram to illustrate the data.

x = time in minutes	number of freshmen
$0 \le x < 10$	101
$10 \le x < 20$	237
$20 \le x < 30$	169
$30 \le x < 40$	79
$40 \le x < 50$	51
$50 \le x < 60$	63
	n =

Solution:

The first step in the answer is to add up the numbers in the right column to find *n*. Here the numbers add up to 700. Thus $n = 700$. To draw the histogram, you place the numbers in the left column along the bottom of the horizontal axis (*x*-axis) and write a label below that saying that it represents time in minutes. You now label the vertical axis (*y*-axis) with the word frequency or the title of the right column. You also need to determine a scale for the vertical axis. Since by numbers go from 51 to 237, going by 50s would make sense. We could also choose to go by 20s, but that would require several more tick marks along the axis. Remember that the labels on the horizontal axis go at the beginning and ending of the bar when the data is grouped (given in intervals). The histogram would look like the one below.



6. The frequency distribution below summaries the hourly wages of the workers at ASU food service. Complete the frequency distribution and draw a histogram to illustrate the data.

x = hourly wage	number of employees
$\$4.00 \le x < \$5.50$	21
$\$5.50 \le x < \$7.00$	35
$\$7.00 \le x < \$8.50$	42
$\$8.50 \le x < \$10.00$	27
$\$10.00 \le x < \$11.50$	18
$\$11.50 \le x < \$13.00$	9
	n =

Solution:

The first step in the answer is to add up the numbers in the right column to find n. Here the numbers add up to 152. Thus $n = 152$. To draw the histogram, you place the numbers in the left column along the bottom of the horizontal axis (x-axis) and write a label below that saying that it represents hourly wage. You now label the vertical axis (y-axis) with the word frequency or the title of the right column. You

also need to determine a scale for the vertical axis. Since by numbers go from 9 to 42, going by 10s would make sense. Remember that the labels on the horizontal axis go at the beginning and ending of the bar when the data is grouped (given in intervals). The histogram would look like the one below.