Math 102
Exam 4
04/27/2009

Name:____

1. A researcher wants to research public opinion on the Minnesota State Senate's proposal to raise taxes on people who make more than $100,000 a year. She goes to the Minikahda Country Club and asks members to fill out a survey. Although some people refuse to answer, she eventually gets 50 responses. Of those that responded, 8 say that they are willing to pay higher taxes, 29 say that they'd be unhappy paying higher taxes, but would live with it, and 13 say that they would move to another state if the proposal is adopted.

   (a) (3 points) What is the population in this survey?

   The population is all people living in Minnesota (the stated goal of the researcher is to understand public opinion on this proposal)

   (b) (3 points) What is the sample in this survey?

   The sample is the 50 people who were surveyed (I also accepted "all people who were surveyed including those who refuesed to answer")

   (c) (4 points) What forms of bias, if any, do you see in the way data was collected in this survey? Explain your reasoning.

   This survey has clear selection bias. The sample is not a random sample. Only members of a particular country club who were present at a partucular time were included in the survey.

   There is also some evidence of a non-respone bias, since we are told that some people refused to answer. To decide this for sure, we would need to look at the proportion of people asked who actually agreed to take the survey.

   (d) (4 points) Based on this study, what conclusions, if any, can be reached about public opinion on the new tax proposal? Explain your reasoning.

   We might possibly be able to reach a conclusion about how members of this particular country club feel about the proposed tax increase, but beyond that, since the sample is so flawed, we learn nothing about public opinion about this proposal in general.

2. A student has scored 82, 76, and 74 on the first three exams in a History class. Suppose that a fourth exam will be given next week and that the student's final grade will be computed based solely on their scores on these four exams (all four are out of 100 points)

   (a) (4 points) What score would the student need to get on this exam in order for their final average to be 80?

   Let x be the missing score. To get an 80 average, we need to have $\frac{82+76+74+x}{4} = 80$, or $82+76+74+x = 320$. Therefore, we must have $232 + x = 320$, or $x = 88$.

   (b) (4 points) What score would the student need to get on this exam in order for their final average to be 90?

   Let x be the missing score. To get an 90 average, we need to have $\frac{82+76+74+x}{4} = 90$, or $82+76+74+x = 360$. Therefore, we must have $232 + x = 360$, or $x = 128$.

   Since there are only 100 points available on the exam, getting up to a 90 average would be impossible.

   (c) (4 points) Supposing that they scored 88 on the fourth exam, would the student rather have their final grade computed using the median or the mean?

   The mean of the 4 scores is 80, while the median is 79, so the student would prefer that the mean was used.

3. The following data set gives the daily high temperature in Fargo, ND from April 12th to April 24th, 2009 (source: www.wunderground.com) $\{55, 58, 62, 66, 63, 57, 44, 52, 50, 61, 58, 80, 46\}$

(a) (6 points) Make a stem and leaf display for this data set.

4	4 6
5	0 2 5 7 8 8
6	1 2 3 6
7	
8	0

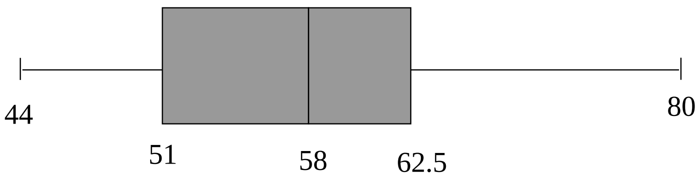
(b) (6 points) Find the mean and midrange of this data set.

mean: $\overline{x} = \dfrac{752}{13} = 57.8°F$

midrange: $\dfrac{44+80}{2} = 62°F$

(c) (8 points) Find the 5 number summary for this data set and draw a “Box-and-Whisker” plot.

min: 44, Q_1: 51, Q_2: 58, Q_3: 62.5, Max: 80



(d) (5 points) Which measure of center do you think best describes the “middle” of this data set? Explain your reasoning.

Notice that this data set has a positive outlier (80). Because of this, we would expect that the median would be the best measure of center. In practice, the mean is $57.8°F$, the median and mode are both $58°F$, and the midrange is $62°F$. The positive outlier seems to be effecting the midrange, so it is not the best measure. There is only one repeated value and the data is numerical, so the mode is not particularly useful. Since the median is less than the mean, the outlier is not effecting the mean, so both the mean and median end up being good measures of center for this data set.

4. Given the following frequency table:

x	Freq.	$x \cdot f$	$x - \overline{x}$	$(x-\overline{x})^2$	$(x-\overline{x})^2 \cdot f$
4	2	8	-5	25	50
7	6	42	-2	4	24
8	5	40	-1	1	5
10	2	20	1	1	2
12	4	48	3	9	36
22	1	22	13	169	169
Total:	20	180			286

(a) (8 points) Compute the mean and median of the data in this table.

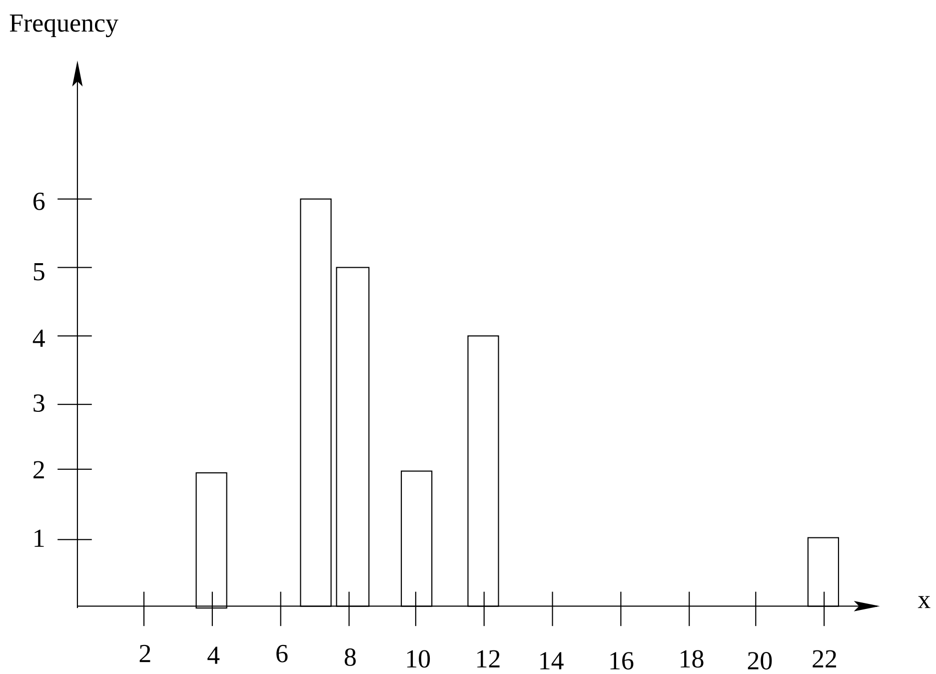
mean: $\overline{x} = \frac{180}{20} = 9$

median: (we average the 10th and 11th data values) $\frac{8+8}{2} = 8$

(b) (10 points) Compute the standard deviation of the data by completing the table above.

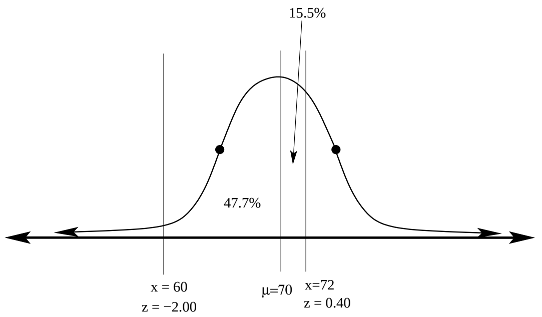
standard deviation: $s = \sqrt{\frac{286}{19}} \approx 3.88$

(c) (6 points) In the space provided, make a frequency histogram for the data in the table above. Be sure to label your axes.



5. (5 points each) Suppose that the height of people in that Fargo metropolitan area is normally distributed with a mean of 70 inches and a standard deviation of 5 inches. Also suppose that the total population of the Fargo metropolitan area is 175,000.

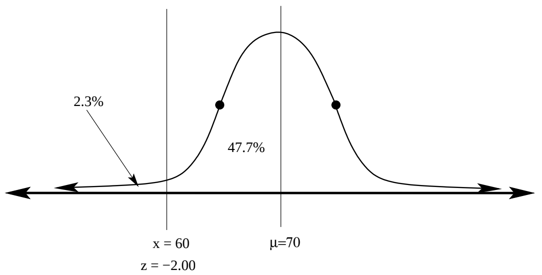
(a) What *percentage* of the population is *between* 5 and 6 feet tall (60 and 72 inches)?



$$z_1 = \frac{60-70}{5} = -2.00,\ z_2 = \frac{72-70}{5} = 0.40$$

Therefore, using the z-table and the digram above, $A = A_1 + A_2 = 0.477 + 0.155 = 0.632$, or 63.2%.

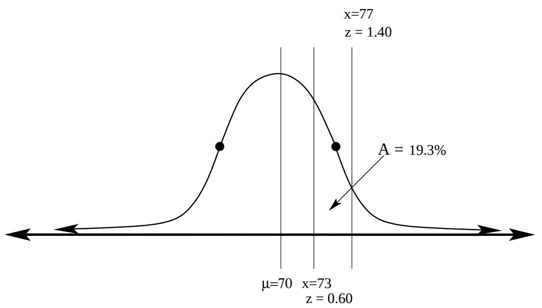
(b) What *percentage* of the population is *under* 5 feet (60 inches) tall?



$$z = \frac{60-70}{5} = -2.00, \text{ so, using the } z\text{-table, } A = 0.477.$$

However, we want the percentage of the population *below* 60 inches in height, so we compute $50 - 47.7 = 2.3\%$.

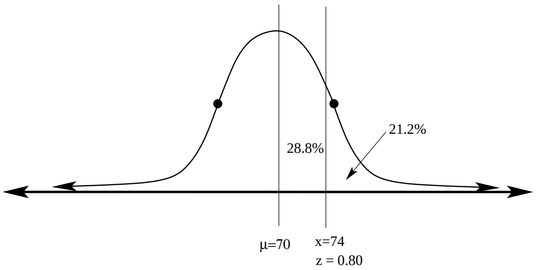
(c) What *percentage* of people are *between* 73 and 77 inches tall?



$$z_1 = \frac{73-70}{5} = 0.60,\ z_2 = \frac{77-70}{5} = 1.40, \text{ so } A_1 = .226 \text{ and } A_2 = .419$$

Therefore, using the z-table and the digram above, $A = A_2 - A_1 = 0.419 - 0.226 = 0.193$, or 19.3%.

(d) **How many people** are *over* 74 inches tall?

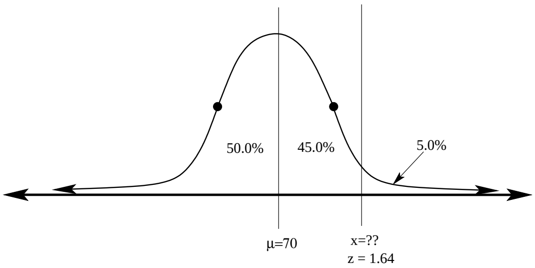


$z = \dfrac{74-70}{5} = 0.80$, so, using the z-table, $A = 0.288$.

To find the percentage of the population *above* 74 inches in height, we compute $50 - 28.8 = 21.2\%$.

However, we want the *number of people* that satisfy this description, so we compute $(175,000)(.212) = 37,100$. We conclude that 37,100 people in the Fargo metro area are over 74 inches tall.

(e) **How tall** would a person need to be to be *taller than* 95% of the people in the Fargo metro area?



We want to find the x-score so that 95% of the population is below this score. We see from the z-table that when $A = 0.950$, $z \approx 1.64$.

Therefore, to find our raw score, $x = (z)(\sigma) + \mu = (1.64)(5) + 70 = 78.2$ inches.

Hence a person must be 78.2 inches tall in order to be taller than 95% of the popoulation in the Fargo metro area.

Extra Credit: (5 points) A company with 100 employees currently has a mean salary of $40,000 per year. Suppose that 10 employees, each of whom make $50,000 per year are laid off. What is the new mean salary of the company after these layoffs?

To compute the new mean salary, we subtract the salary of the laid off employees from the total of the salaries at the company and then divide by 90, the new total number of employees.

Then $\overline{x} = \dfrac{100(\$40,000) - 10(\$50,000)}{90} \approx \$38,888.89$