MATH 1203 – Practice Exam 1

This is a practice exam. The actual exam may be different from this and contains fewer questions

1. Please state, in your own words, what the following terms mean:

   Population, sample, random sample, numeric variable, categorical variable, ordinal and nominal variable, homogeneous and heterogeneous distribution, normal distribution, skewed distribution, mean, mode, median, range, variance, standard deviation, Q1, Q2, IQR, percentile, empirical rule, estimates for standard deviation, outlier, *and any other term we discussed in class.*

2. Please decide if the following statements are true or false.
   a) To compute the variance, you must first compute the median
   b) A box plot can help you to decide whether a distribution is skewed to the right or left.
   c) The mean is influenced by extremely large or extremely small values
   d) The standard deviation is the square of the variance
   e) The variance is the square of the standard deviation
   f) To compute the *variance*, you must first compute the *mode*.
   g) To draw a box plot, you need the *highest* and *lowest* values as well as the *mean* and the *standard deviation*.
   h) To compute the median and the quartiles, you need to sort the data first
   i) If your SAT scores are the "80th percentile", it means that 80% of SAT scores are bigger than your score.
   j) For a bell-shaped distribution, approximately 95% of all data values are between $\overline{x} - 2s$ and $\overline{x} + 2s$
   k) For a bell-shaped distribution, the standard deviation is approximately equal to ½ * IQR
   l) If the variance for a distribution is zero, all data values are the same.
   m) If the variance for a distribution is negative, the data values are spread far away around the mean
   n) If a distribution is skewed to the left, the mean is larger than the median

3. Below is a short segment from a (fictitious) survey questionnaire. How many variables can you identify? For each variable, state whether it is nominal, ordinal, or numeric.

   ```
   Please state your height in cm:
         _____

   What is the color of your eyes:
         [ ] green
         [ ] brown
         [ ] other

   In which age group do you fall:
      [ ] under 20 years
      [ ] 20 to 50 years
      [ ] over 50 years
   ```

4. For which type of variable can you compute the mean? How about the mode? The median? Variance? Q1 and Q3? Percentiles?

5. Use the frequency distribution listed below to answer the following questions:

OPINION OF FAMILY INCOME

		Freq.	Percent	Valid Percent	Cumulative Percent
Valid	FAR BELOW AVERAGE	177	6.1	6.2	6.2
	BELOW AVERAGE	700	24.1	24.4	30.6
	AVERAGE	1384	47.7	48.3	78.8
	ABOVE AVERAGE	534	18.4	18.6	97.5
	FAR ABOVE AVERAGE	73	2.5	2.5	100.0
	Total	2868	98.8	100.0	
Missing		36	1.2		
Total		2904	100.0		

a) What percentage of the sample thinks their family income is far below average?

b) What percentage of the sample thinks their family income is less than average

c) What percentage of the sample thinks their family income is average or better

6. Use the frequency distribution listed below to answer the following questions:

OPINION OF NUCLEAR POWER

		Freq.	Percent	Valid Percent	Cumulative Percent
Valid	Strongly dislike	30	30	30	30
	Dislike	20	20	20	50
	Neutral				
	Like	30	30	30	
	Strongly like	10	10	10	100.0
	Total	100	100	100.0	
Missing		36	1.2		
Total		2904	100.0		

a) What is the missing percentage in the "valid percent" column?

b) What is the missing percentage in the "cumulative percent" column?

7. Suppose data from a random sample of cars shows the following miles per gallon figures and main colors for a variety of cars:

Mpg's: 12, 23, 32, 33, 43, 15, 21, 22, 44, 51, 44, 32, 27, 26, 23, 17, 29, 38, 28, 28, 20, 21,30, 31, 32, 33, 41, 39, 25, 24, 26, 25
Colors: white, black, blue, red, white, white, blue, red, yellow, black, white, blue, brown, yellow, red, black, white, black, white, black, brown, yellow, red, black, white, white, white, black, red, brown, blue, white, blue

Create a histogram (with 4 bins) and a frequency distribution for the variables, as appropriate. Find the mean, median, and mode, as appropriate for the variable. Also find the 80th percentile.

8. Suppose a random sample of size 10 taken from the GSS survey shows that the age for that sample are as follows:

28, 32, 33, 28, 29, 30, 30, 28, 32, 30

a) Find the mode, the mean, and the median for this data
b) Find the range, variance, and standard deviation for this data
b) Find Q_1, Q_3, and the IQR for this data
c) Draw an approximate box plot for this data

9. A random sample of size 25 selected from the GSS shows the following distribution for the highest degree of the respondent.

RS HIGHEST DEGREE

	Freq	Percent	Valid Percent	Cumulative Percent
LT HIGH SCHOOL (0)	10	40.0	40.0	40.0
HIGH SCHOOL (1)	9	36.0	36.0	76.0
JUNIOR COLLEGE (2)	1	4.0	4.0	80.0
BACHELOR (3)	3	12.0	12.0	92.0
GRADUATE (4)	2	8.0	8.0	100.0
Total	25	100.0	100.0	

a) Is this distribution heterogeneous or homogeneous?
b) Find the mode and median for this distribution.
c) Using the integer codes, find the mean.
c) Find Q_1 and Q_3

10. Please match the following statements to the distribution pictures below.

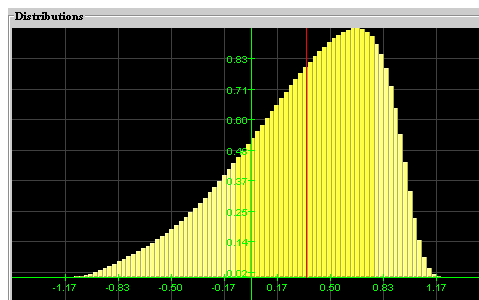i) Skewed to the Right   ii) Skewed to the Left
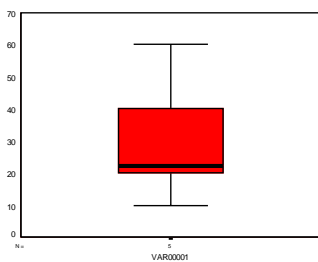
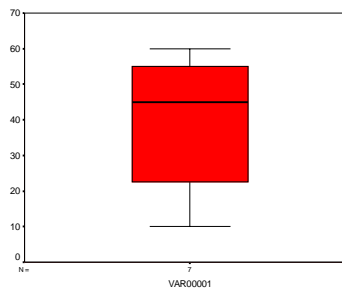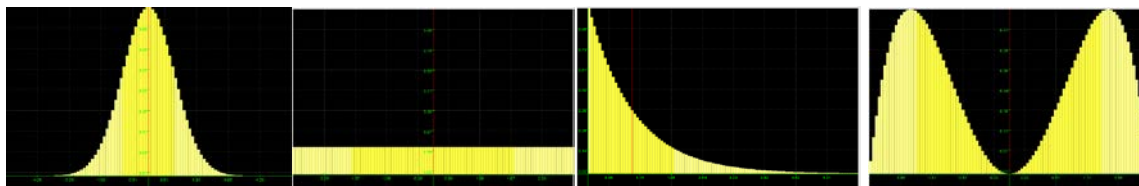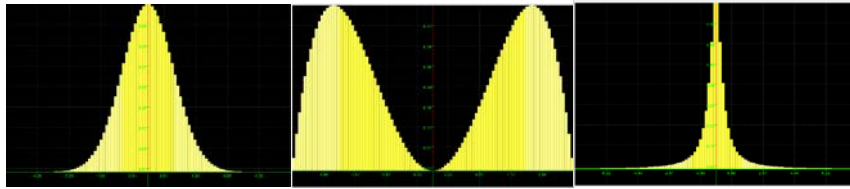(a) Mean less than median   (b) median less than mean
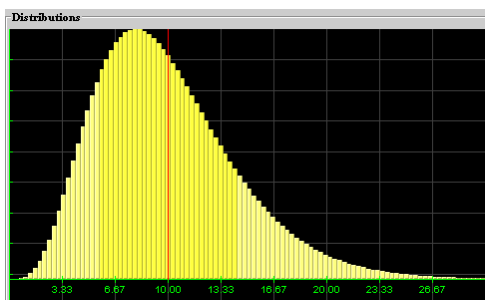
(1)

(2)

(*)

(**)

11. Which of the following distributions is homogeneous, which is heterogeneous?

12. Below are three distributions, all with mean zero. Which has the largest, which the smallest standard distribution?

13. Suppose a normal (bell-shaped) distribution has mean 10, with a minimum value of 2 and a maximum value of 18. Provide a good estimate for the standard deviation and provide a reason for your assessment.

14. Suppose a normal (bell-shaped) distribution has $Q_1 = 20$ and $Q3 = 50$. What would be a good estimate for the standard deviation?

15. Suppose a box plot for the number of children in our GSS survey looks as follows:

NUMBER OF CHILDREN
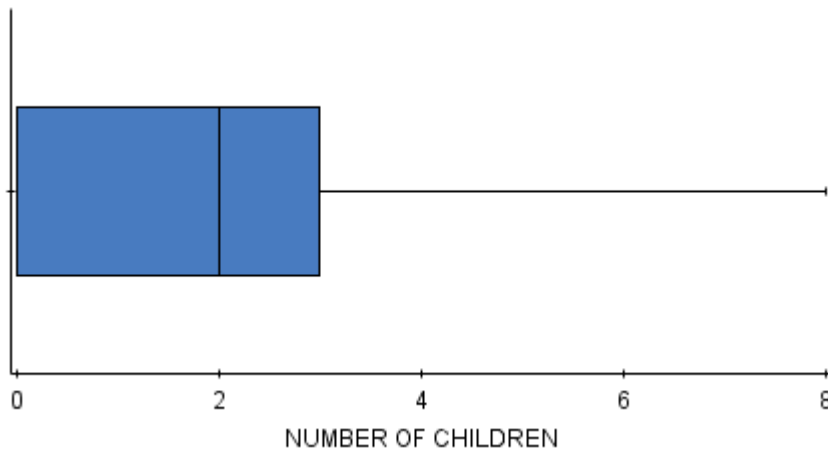
- What is the lower and upper quartile?
- What is the median?
- What, approximately, might be the mean?
- Estimate the standard deviation from the range
- Estimate the standard deviation from the IQR
- Which values would you consider to be outliers?

16. Suppose a distribution looks as follows:
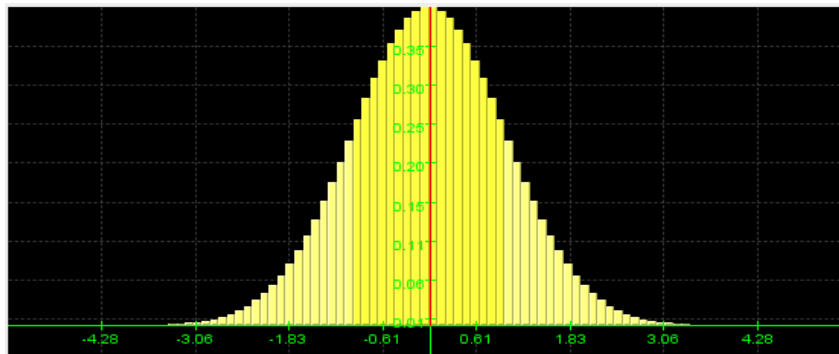
Provide estimates for the mean and the standard deviation and justify your guess.

17. Consider the frequency distribution below, taken from our GSS survey

Frequency table results for OPINION OF FAMILY INCOME:

OPINION OF FAMILY INCOME	Frequency	Relative Frequency
1 - Far below average	145	0.0725
2 - Below average	545	0.2725
3 - Average	885	0.4425
4 - Above average	385	0.1925
5 - Far above average	40	0.02

Is this distribution homogeneous or heterogeneous?
Find Q1, Q3, and the median
Is this distribution approximately normal or skewed? Justify your answer.
Find the 90th percentile

18. Consider the frequency distribution below, taken from our GSS survey:

Frequency table results for LABOR FORCE STATUS:

LABOR FORCE STATUS	Frequency	Relative Frequency
1 - Working full time	1003	0.49628896
2 - Working part time	211	0.104403764
3 - Temp not working	53	0.026224641
4 - Unempl., laid off	74	0.036615536
5 - Retired	336	0.16625433
6 - School	57	0.02820386
7 - Keeping house	227	0.11232063
8 - Other	60	0.029688273

Would it make sense to find Q1, Q3, and the median? Explain.

19. Some question about sample and/or random sample selection. Please check back later for an updated version of the practice exam.

Additional questions are possible. Please review our lecture notes and all quizzes.