MT 143, Exam #1 Review, Spring 2020

Major topics:
- Definitions
- Sampling Methods
- Frequency Distributions
- Making and Interpreting Graphs
- Calculating mean, median and mode, standard deviation, 5-number summary, percentiles, z-score
- Counting methods
- Probability rules
- Mean/expected value and standard deviation of discrete distributions
- Binomial probability distribution

Sample problems.
1. Classify the following variables:

Variable	Categorical	Quantitative	Discrete	Continuous	Nominal	Ordinal	Interval	Ratio
Test Grade (Letter)								
Continent								
Month								
Volume (Sound)								
Cost ($)								
Vegetable								
Goal								
Altitude								
Growth Rate								
Sexual Orientation								
Year of Birth								

2. What does it mean if we say a value represents the 4th percentile?

3. With symmetric/bell-shaped distributions, approximately what percent of the observations are within three standard deviations of the mean?

4. Expressed in percentiles, what does the first quartile represent?

5. Explain the difference between a classical (theoretical) probability and an experimental (observational) probability.

6. A particular model of lie detector test has a 98% probability of correctly detecting someone who is lying, and 99.5% probability of correctly detecting someone who is not lying. A human police officer interviews suspects using the lie detector, hoping to catch the 1 of the 20 suspects worth investigating further who they believe to be lying. If someone tests positively on the lie detector

for lying, what is the probability that the person is actually lying? Construct a tree diagram to model the situation. Should the police continue to use the lie detector, and if so, is it safe to assume the person is definitely guilty?

7. Suppose that Comdell Computer receives its hard drives from Diskco. On average, 3% of all hard disk drives received by Comdell are defective.
   a. Comdell has adopted the following policy: It samples 50 hard drives in each shipment and accepts the shipment if all hard drives in the shipment are not defective. What fraction of shipments will Comdell accept?

   b. Suppose instead that the shipment is accepted if at most one hard drive in the sample is defective. What fraction of shipment will Comdell accept?

   c. What is the probability that a sample size of 50 will contain at least 5 defectives?

8. A researcher is interested in finding out about the buying habits of the American public. She decides to conduct a survey of 50 shoppers at a local mall and asks them how much money they are planning to spend on gifts for an upcoming holiday.
   a. What is the population of interest in this study?

   b. What is the sample used in this study?

   c. What is the parameter being studied?

   d. What statistic is the researcher likely to use to estimate the parameter?

9. A sample of retirees were asked their age. The results are shown below in a table.

Ages of Retirees				
79	80	86	72	76
50	81	71	96	81
88	78	72	65	95
74	105	55	83	63
56	68	65	90	69
58	98	64	84	91
72	75	94	91	79
56	82	90	85	87
99	72	96	76	109
83	81	89	67	69
100	89	84	101	92
63	87	85	87	66

   a. Use the data in the table to construct a stemplot of the data.

   b. The mean, median and mode

   c. The standard deviation and range

d. Calculate the five-number summary for this data.

e. Use that information to construct a simple box plot. Be sure your graph is to scale.

f. Based on this information, what percentile is a 72-year-old retiree?

10. Sketch the Empirical Rule on the graph below. Label the graph clearly.
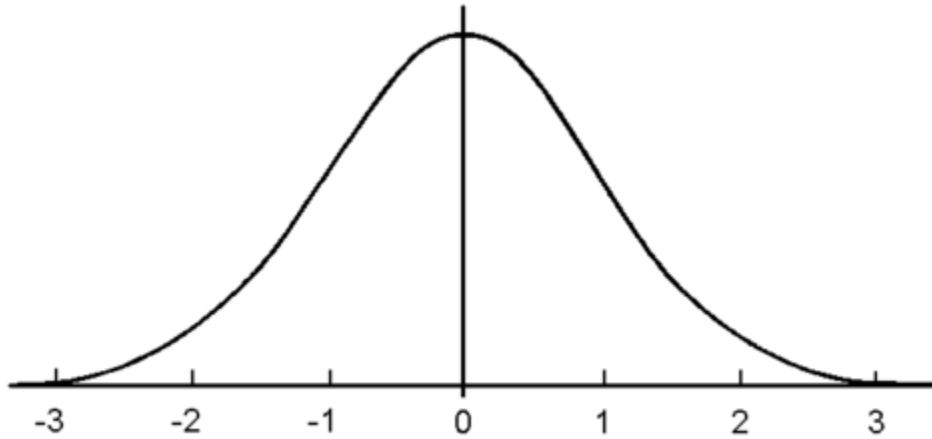
11. If the mean of a normal distribution is 72 points with a standard deviation of 5 points, find the following:
a. The z-score of 88 points.

b. What percentile 88 points represents.

c. The area under the normal curve between 67 and 82 using the Empirical Rule.

12. For each of the distributions shown below, determine if the distribution is roughly symmetric, left skewed, right skewed, roughly uniform, or none of these.

a.

c.

Histogram of "Next Stop Westchester"

b.

Number of Samples vs
10000 Random Real Numbers between 0 and 100
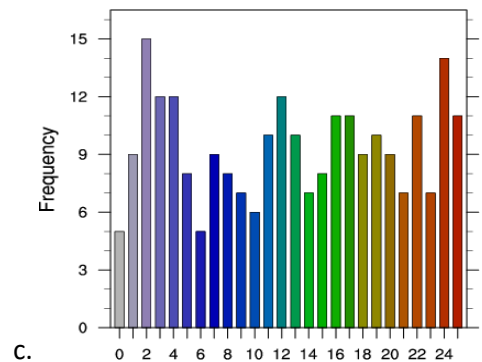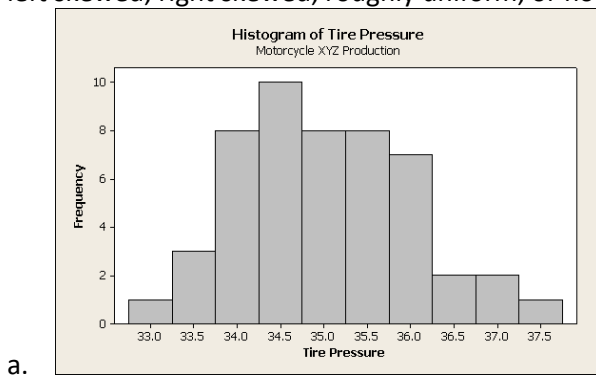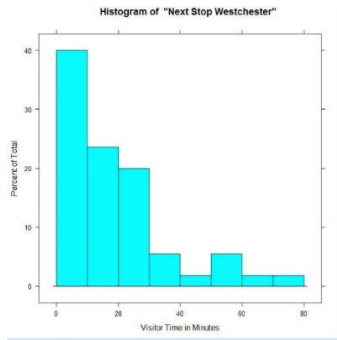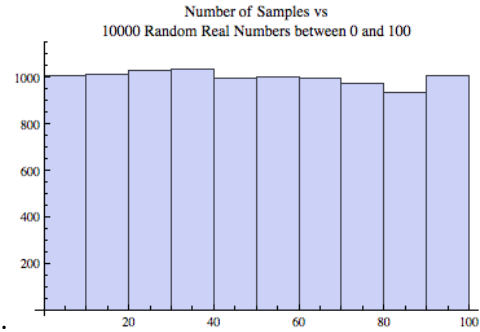
d.

13. The 100 tiles in Scrabble are distributed as follows:

Tile	Number	Tile	Number	Tile	Number	Tile	Number
Blank	2	A	9	B	2	C	2
D	4	E	12	F	2	G	3
H	2	I	9	J	1	K	1
L	4	M	2	N	6	O	8
P	2	Q	1	R	6	S	4
T	6	U	4	V	2	W	2
X	1	Y	2	Z	1		

    a. What is the probability of selecting an R as the first tile in a game?

    b. What is the probability of selecting a vowel (not Y) as the first tile?

    c. What is the probability of not selecting a vowel?

    d. What is the probability of selecting the word BOX in order from the first three tiles in a game?

14. Use the following table to calculate the probabilities requested.

		STUDENT'S CHOICE		
		Art degree	Science degree	Total
GROUP	Boys	25	50	75
	Girls	55	20	75
	Total	80	70	150

    a. What is the probability of a randomly selected person from this study is a girl?

b. What is the probability of a randomly selected person from this study being a girl and wants a science degree?

c. What is the probability of a randomly selected person from this study being a girl or wanting a science degree?

d. What is the probability of being a girl given that the person wants a science degree?

e. Are the variables gender and degree choice independent? Why or why not? Show calculations to justify your answer.

15. Determine the **number** of outcomes in each of the following scenarios.
   a. A local area network requires eight characters for a username and is not case sensitive, but the character can use numbers in the last three digits. How many usernames of this type are there?

   b. Suppose that a lottery has 42 balls, and someone needs 6 matches in any order to win the top prize. How many possible winning number combinations are possible?

   c. Suppose that ten horses are in a particular race. How many ways can the top three horses finish?

   d. Tim is planning to create a word scramble puzzle out of the word ENCYCLOPEDIA. How many different sequences of the letters are possible?

16. Find the probability of each of the following scenarios.
   a. Social Security numbers are composed of 9 digits, and each digit can be any number from 0-9. What is the probability that a randomly selected Social Security number will be all even numbers?

   b. What is the probability of getting a three-of-a-kind in a 5-card poker hand?

17. Evaluate the following expressions.
   a. $\binom{9}{4}$ b. $_3^7P$ c. $_5^{12}C$ d. $0!$

18. Use the following table to find the probability of getting the sum of 14 on rolling two 8-sided dice.

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

5						
6						
7						
8						

19. A certain game is played by drawing cards with numbers on them according to the following distribution.

Score, x	Probability, P(x)
0	0.07
1	0.13
2	0.18
3	0.30
4	0.22
5	0.08
6	0.02
	$\Sigma P(x) = 1.00$

Find the expected value (mean) and the standard deviation of this card game.

20. A charity sells 450 tickets for a raffle, costing $10 per ticket. The top prize is $500, a second prize of $100, and two third prizes of $50. For someone purchasing a ticket, what is the expected value? Interpret the value in the context of the problem.

Value				
Probability				

21. Two thirds of drivers put their seat belt on when they get into a driving simulator. Use that fact to answer the following questions.
    a. If 10 people get into the simulator, what is the probability that exactly 8 of the people will put on their seat belts?

    b. If 10 people get into the simulator, what is the probability that no more than 8 people will put on their seat belts?

    c. What is the expected number of those 10 people who put on their seat belts?

22. The SAT has a mean score of 1498 and a standard deviation of 199.
    a. What is the z-score of 1640?

b. The mean score on the ACT is 21 with a standard deviation of 5.2. Which student scored higher: Abby with a score of 28 on the ACT, or Barbara with a score of 1910 on the SAT?

23. A particular church choir was asked their ages. The results are shown below in a stemplot below. Key: 3|1 = 31.

For the data shown, find the following statistics:
   a. The mean, median and mode

   b. The standard deviation and range

   c. Calculate the five-number summary for this data.

   d. Use that information to construct a simple box plot. Be sure your graph is to scale.

Stemplot of ages in church choir

```
1 |5 5 6
2 |0 3 3 4 6 6 8
3 |1 4 6
4 |3 3
```

24. Find graphs and statistics of data. Use the labs and quizzes as additional examples.