Name: Key STAT 210 Practice Final Exam

1) Explain how the variable "test grade" could be measures as both a categorical variable and a quantitative variable. (2 pts)

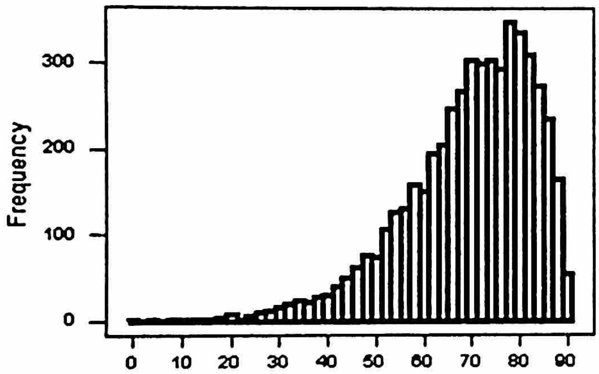
Q: %'s or Raw Score C: A, B, C, D, etc.

2) If you are dealing with quantitative data what are two ways to graphically represent your data? (2 pts)

Histogram, Boxplot, Stemplot, Dotplot

3) Based on the histogram shown answer (T/F) for each statement below: (1 pt. each)

Ske



a) The mean is less than the median. T

b) The Standard Deviation is the best measure of spread. F (skewed)

c) The distribution is symmetrical. F (skewed left)

d) The data represented is categorical data. F (quantitative)

e) The graph is a histogram. T

4) If the mean score on a test was a 71% and the standard deviation was 8%, what would be the approximate score of someone in the 80th percentile? (4 pts)

look up .8000 to find z-score of ≈ .84

$Z = \frac{X - \text{mean}}{SD}$ $8(.84) = X - 71$

$.84 = \frac{X-71}{8}$ $\boxed{X = 77.72}$

5) If the probability of winning on a scratch off ticket is 8% and you decide to buy some tickets. (4pts each) (MINITAB)

a) What is the probability that you win on 2 out of 15 tickets?

* SKIP *

b) What is the probability that if you purchase 50 scratch-off tickets you will win on more than 5 of them?

* SKIP *

6) What is the probability that if you randomly select a letter from the word "ST<u>A</u>T<u>I</u>ST<u>I</u>CS" and then roll a 6-sided die, that you get a vowel and a number <u>greater</u> than 3? (3 pts)

$$\left(\frac{3}{10}\right)\left(\frac{3}{6}\right) = \frac{9}{60} = \boxed{\frac{3}{20}} \text{ or } \boxed{0.15}$$

7) You got an 82% on a math test where the class average was a 77% and the standard deviation was a 3%, and then you got a 79% on a science test where the class average was a 71% with a standard deviation of 4%.

a) Using Z-scores which class did you do better in compared to the rest of the class? (3 pts)

M

$$z = \frac{82-77}{3}$$

$$z = 1.67$$

S

$$z = \frac{79-71}{4}$$

$$z = 2$$

b) What was your percentile ranking in each class? (2pts)

M (95.25%) S (97.72%)

c) In your math class approximately what percent of people scored between a 70% and an 82%? (3 pts)

$$z = \frac{70-77}{3} = -2.33 \qquad z = \frac{82-77}{3} = 1.67$$

(.0099) (.9525)

.9525 − .0099

$$.9426 \approx \boxed{94.26\%}$$

8) Given the summary of a data set gives the following values:
Mean = 23, Median = 30, Min = 9, Q1 = 24, Q3 = 34, Max = 40

a) Show if any numbers are outliers (3 pts)

$IQR = Q_3 - Q_1 = 10$ $\quad Q_1 - 15 = 9$
$1.5(IQR) = 1.5(10) = 15$ $\quad Q_3 + 15 = 49$ } All #'s between

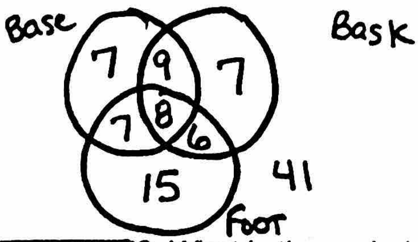
NO OUTLIERS!

b) Would the distribution be skewed left, right, or symmetric based on the values you are given above? (Support your answer) (2 pts)

Skewed left $\quad$ mean < median

9) At a school there are 100 students in the Senior Class and:
8 Students play Baseball, Basketball and Football
15 Students play Baseball and Football
17 Students play Baseball and Basketball
14 Students play Basketball and Football
31 Students play Baseball
36 Students play Football
30 Students play Basketball

- Draw a Venn diagram to represent this scenario, make sure to include students that do not play any sports. (3 pts)



- ~~What is the probability someone plays football? What is the probability that someone plays football given they also play baseball?~~ (2 pts) *(SKIP)*

- What is the probability that you randomly select one student and they do not play any of these sports? (2 pts)

$\frac{41}{100} = .41$

- What is the probability that if you select two students, they both play *all three* sports? (2 pts)

$\frac{8}{100} \cdot \frac{7}{99} = \frac{56}{9900} \approx .0057$

FOR PROBLEMS 10-11 CONDUCT THE APPROPRIATE HYPOTHESIS TEST AND MAKE SURE TO INCLUDE ALL NECCESARRY PARTS OF THE TEST (10 pts each)

10) 100 Randomly selected Americans were asked how many hours of T.V they watch each week, and the 100 Americans had a mean number of hours of 11.3 hours and a standard deviation of 2.9 hours. A professor believes that people in America watch more T.V than people in Canada. So, he also asked 110 randomly selected Canadian citizens the same question and found out that the mean of the 110 Canadians was 9.9 hours with a standard deviation of 3.6 hours. Is there enough evidence to support the professor's claim?

$H_0: \mu_A = \mu_C$ Population A: All Americans Parameter: μ_A & μ_C Avg # of
$H_a: \mu_A > \mu_C$ Population C: All Canadians hours a week watching T.V

2-Sample T-test $n_A = 100$ t-score ≈ 3.12 d.f = 205

Conditions:

$\bar{X}_A = 11.3$ P-value ≈ .001

Random Samp A & Random Samp C

$S_A = 2.9$ Since the p-value is less than 5% we reject Ho, which means

$10n_A < \text{Pop } A$ & $10n_C < \text{Pop } C$ $n_C = 110$

$\bar{X}_C = 9.9$ the evidence suggests people in

$n_A \geq 30$ & $n_C \geq 30$

$S_C = 3.6$ America do watch more T.V than people in Canada.

11) An employee did not believe a companies claim that the average salary at the company was $57,000. The employee randomly selected 50 employees at the company and found out that the mean of the sample was $59,200 with a standard deviation of $13,500. Is this enough statistical evidence that the company was incorrect in their statement?

Population: All people at a company Parameter: μ = average salary in $

1-Sample t test

$n = 50$

$\bar{X} = 59{,}200$

$S = 13{,}500$

Conditions:
1) Random Samp.
2) 10n < Pop Size
3) $n \geq 30$

t-value = 1.15
d.f = 49
p-value ≈ .255

Since the p-value is greater than 5% we reject the Ho, which means that we do not have enough evidence to suggest (Ha) the average salary at the company is not $57,000.

$H_0: \mu = 57000$

$H_a: \mu \neq 57{,}000$

12) Sara is taking a test in her science class and her math class, if the class average in science was a 72% and the standard deviation was 8%, compared to an average of 78% and a standard deviation of 2% in her math class. Which tests did she do better on compared to the rest of the class given her score on the science test was a 90% and the math test score was an 83%? (Explain)

S: $z = \frac{90-72}{8} = 2.25$

M: $z = \frac{83-78}{2} = 2.5$

Better in math, higher z-score

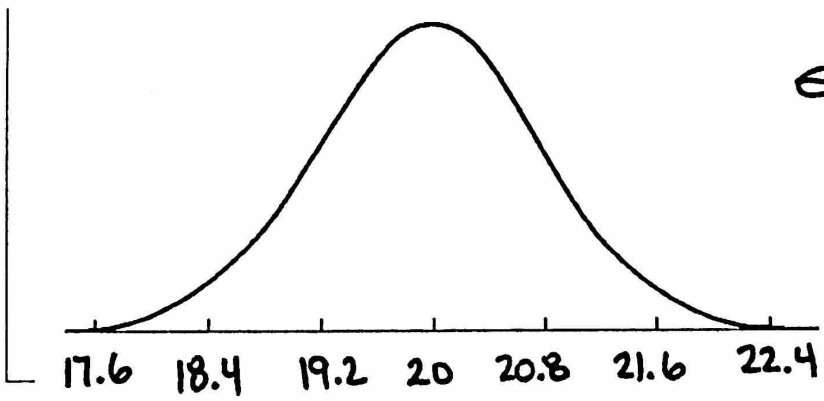
13) Read the sentences below and then identify each **underlined (bold)** number by labeling it with the proper statistical symbol. (6 pts)

Jordan wanted to figure out the average number of people living in each house in Maryland. He randomly selected **100** houses from the phone book and called each one to ask how many people live in the house. The average number he calculated for the people that answered the phone was **4.8.** He knows based off a previous study the standard deviation for the number of people per household in Maryland is **1.7**. In another study he asked those same 100 people if they enjoyed living in their neighborhood and **73%** said that they did.

$n = 100$ $\quad \sigma = 1.7$

$\bar{x} = 4.8$ $\quad \hat{p} = .73$

14) If mean of a population that is normally distributed is 20 and the standard deviation of the population is 4. Sketch the sampling distribution of sample means for a sample size of 25. (4 pts)



$\sigma_{\bar{x}} = \frac{4}{\sqrt{25}} = \frac{4}{5} = .8$

15) Lie detectors are based on measuring changes in the nervous system. The assumption is that lying will be reflected in physiological changes that are not under the voluntary control of the individual. When a person is telling the truth, the galvanic skin response scores have a distribution that is normal with a mean of 51.6 and a standard deviation of 9. (Assume ALL Conditions are met)

What is the probability that a random sample of 30 people will have an average score less than 50.9? (6 pts)

$\sigma_{\bar{x}} = \frac{9}{\sqrt{30}} \approx 1.64 \qquad Z = \frac{50.9 - 51.6}{1.64} \approx -.43 \qquad (.3336)$

16) If someone believes the average temperature in a specific state is over 81° and they take a sample of 40 days and record the temperatures for that state. The average temperature for the 40 days is 83°. They conduct an appropriate hypothesis test and discover the p-value is equal to 0.04. Explain in detail what the p-value is saying in the context of this problem.

There is a 4% chance the average temperture in the state would be 83° or higher over a random sample of 40 days, if the actual average temp. in the state is 81°

17) A publishing company pays its sales staff $600 a week plus commission of $0.50 per book sold. For example, a salesman who sold 440 books last week earned $600 + $0.50(440) = $820

Statistic	Books Sold	$ Earned
Mean	640 (.5) 320	920
Standard Deviation	360 (.5) 180	180
IQR	450 (.5) 225	225
Maximum	1420 (.5) 710	1310

a. The table above shows summary statistics for the number of books the large sales staff sold last week. Fill in the table above to show the statistics for the pay these people earned.

b. Using the mean and standard deviation calculated in part A find the Z-score for someone who earned $1170.

$$Z = \frac{1170 - 920}{180} = \frac{250}{180} \approx 1.39$$

18) A roadway construction process uses a machine that pours concrete onto the roadway and measures the thickness of the concrete so the roadway will measure up to the required depth in inches. The concrete thickness needs to be consistent across the road, but the machine isn't perfect, and it is costly to operate. Since there's a safety hazard if the roadway is thinner than the minimum 23-inch thickness, the company sets the machine to average 26 inches for the batches of concrete. They believe the thickness level of the machine's concrete output can be described by a normal model with a standard deviation of 1.75 inches. [Show all work]

a. What percent of the concrete roadway is under the minimum depth?

$$z = \frac{23-26}{1.75} = \frac{-3}{1.75} \approx -1.71 \qquad (.0436)$$

4.36%

b. The company's lawyers insist that no more than 3% of the output be under the limit. Because of the expense of operating the machine, they cannot afford to reset the mean to a higher value. Instead, they try and reduce the standard deviation to achieve the "only 3% under" goal. What standard deviation must they attain to reach the goal?

$z \approx -1.88$ (look up .0300 in chart) $\quad -3 = -1.88(s)$

$$-1.88 = \frac{23-26}{s} \qquad \frac{-1.88}{1} = \frac{-3}{s} \qquad \frac{-3}{-1.88} = \boxed{s \approx 1.6}$$

19) Using the table shown for Hours worked and # of products sold, answer the following three questions.

a) What is the equation for the linear regression line to predict products sold based on the number of hours worked?

$\hat{y} \approx 5.76x + .4897$

Hours	# Sold
2	18
5	22
7	38
9	54
12	72

b) What type of correlation does the data show?

strong positive linear

c) Using your prediction equation, how many products sold would you predict for someone who worked 7 hours?

$\hat{y} \approx 40.81$

d) What is the residual for someone who worked 9 hours?

obs $(y) = 54$

pre $(\hat{y}) = 52.33$

$54 - 52.33$

≈ 1.67