STA 13 - Final Exam Review

Name_____

SHORT ANSWER. Write your answer on a separate sheet of paper.

Provide an appropriate response.

1) Define the terms population, sample, parameter and statistic.

2) Distinguish between categorical and quantitative data. Give an example for each.

3) Define continuous and discrete data and give an example of each.

4) Histograms and Pareto charts are both bar charts. What is the significant difference between the two?

5) We want to compare two different groups of students, students taking Composition 1 in a traditional lecture format and students taking Composition 1 in a distance learning format. We know that the mean score on the research paper is 85 for both groups. What additional information would be provided by knowing the standard deviation?

6) A company advertises an average of 42,000 miles for one of its new tires. In the manufacturing process there is some variation around that average. Would the company want a process that provides a large or a small variance? Justify your answer.

7) Marla scored 85% on her last unit exam in her statistics class. When Marla took the SAT exam, she scored at the 85 percentile in mathematics. Explain the difference in these two scores.

Find the mean and median for each of the two samples, then compare the two sets of results.

8) A comparison is made between summer electric bills of those who have central air and those who have window units.

	May	June	July	Aug	Sept
Central	$32	$64	$80	$90	$65
Window	$15	$84	$99	$120	$40

Solve the problem.

9) The data set below consists of the scores of 15 students on a quiz. For this data set, which measure of variation do you think is more appropriate, the range or the standard deviation? Explain your thinking.

90 90 91 91 89
90 89 91 91 90
60 90 89 90 91

Find the z-score corresponding to the given value and use the z-score to determine whether the value is unusual. Consider a score to be unusual if its z-score is less than -2.00 or greater than 2.00. Round the z-score to the nearest tenth if necessary.

10) A body temperature of 99.5° F given that human body temperatures have a mean of 98.20° F and a standard deviation of 0.62°.

Determine which score corresponds to the higher relative position.

11) Which is better, a score of 92 on a test with a mean of 71 and a standard deviation of 15, or a score of 688 on a test with a mean of 493 and a standard deviation of 150?

Find the indicated measure.

12) The weights (in pounds) of 30 newborn babies are listed below. Find Q_1.

5.5 5.7 5.8 6.0 6.1 6.1 6.3 6.4 6.5 6.6
6.7 6.7 6.7 6.9 7.0 7.0 7.0 7.1 7.2 7.2
7.4 7.5 7.7 7.7 7.8 8.0 8.1 8.1 8.3 8.7

Provide an appropriate response.

13) Suppose that all the values in a data set are converted to z-scores. Which of the statements below is true?

A: The mean of the z-scores will be zero, and the standard deviation of the z-scores will be the same as the standard deviation of the original data values.
B: The mean and standard deviation of the z-scores will be the same as the mean and standard deviation of the original data values.
C: The mean of the z-scores will be 0, and the standard deviation of the z-scores will be 1.
D: The mean and the standard deviation of the z-scores will both be zero.

Find the indicated probability.

14) A die with 8 sides is rolled. What is the probability of rolling a number less than 7?

Find the indicated complement.

15) The probability that Luis will pass his statistics test is 0.58. Find the probability that he will fail his statistics test.

Find the indicated probability.

16) The table below describes the smoking habits of a group of asthma sufferers.

	Nonsmoker	Occasional smoker	Regular smoker	Heavy smoker	Total
Men	376	31	64	41	512
Women	341	46	74	37	498
Total	717	77	138	78	1010

If one of the 1010 people is randomly selected, find the probability that the person is a man or a heavy smoker.

Find the indicated probability. Express your answer as a simplified fraction unless otherwise noted.

17) The following table contains data from a study of two airlines which fly to Small Town, USA.

	Number of flights which were on time	Number of flights which were late
Podunk Airlines	33	6
Upstate Airlines	43	5

If one of the 87 flights is randomly selected, find the probability that the flight selected arrived on time given that it was an Upstate Airlines flight.

Find the indicated probability. Round to the nearest thousandth.

18) A study conducted at a certain college shows that 56% of the school's graduates find a job in their chosen field within a year after graduation. Find the probability that among 6 randomly selected graduates, at least one finds a job in his or her chosen field within a year of graduating.

Solve the problem.

19) A study revealed that 45% of college freshmen are male and that 18% of male freshmen earned college credits while still in high school. Find the probability that a randomly chosen college freshman will be male and have earned college credits while still in high school.

Provide an appropriate response.

20) List the four requirements for a binomial distribution. Describe an experiment which is binomial and discuss how the experiment fits each of the four requirements.

21) Anna uses the letter X to represent the possible sequences of heads and tails that can be obtained when a coin is flipped three times. The possible sequences together with their probabilities are listed below:

X	Probability of sequence X
HHH	1/8
HHT	1/8
HTH	1/8
HTT	1/8
THH	1/8
THT	1/8
TTH	1/8
TTT	1/8

Is X a random variable? Why or why not? If it is not, which associated variable is a random variable? Give the probability distribution of the associated random variable.

Explain why the following is a valid probability distribution. Then, find the mean of the given probability distribution.

22)

x	P(x)
0	0.42
1	0.12
2	0.34
3	0.05
4	0.07

Provide an appropriate response.

23) In a game, you have a 1/36 probability of winning $81 and a 35/36 probability of losing $3. What is your expected value?

Solve the problem.

24) Compute $\begin{pmatrix} 9 \\ 4 \end{pmatrix}$.

25) It a recent study of college students indicated that 30% of all college students had at least one tattoo. A small private college decided to randomly and independently sample 15 of their students and ask if they have a tattoo. Find the probability that exactly 5 of the students reported that they did have at least one tattoo.

Find the indicated probability. Round to three decimal places.
26) Find the probability of at least 2 girls in 8 births. Assume that male and female births are equally likely and that the births are independent events.

Find the indicated value.
27) $z_{0.02}$

Assume that the weight loss for the first month of a diet program varies between 6 pounds and 12 pounds, and is spread evenly over the range of possibilities, so that there is a uniform distribution. Find the probability of the given range of pounds lost.
28) More than 10 pounds

Find the indicated probability.
29) The weekly salaries of teachers in one state are normally distributed with a mean of $490 and a standard deviation of $45. What is the probability that a randomly selected teacher earns more than $525 a week?

Solve the problem. Round to the nearest tenth unless indicated otherwise.
30) Human body temperatures are normally distributed with a mean of 98.20°F and a standard deviation of 0.62°F. Find the temperature that separates the top 7% from the bottom 93%. Round to the nearest hundredth of a degree.

Provide an appropriate response.
31) SAT verbal scores are normally distributed with a mean of 430 and a standard deviation of 120 (based on data from the College Board ATP). (a) If a single student is randomly selected, find the probability that the they score above 500. (b) If a sample of 35 students are selected randomly, find the probability that the sample mean is above 500. These two problems appear to be very similar. Which problem requires the application of the central limit theorem, and in what way does the solution process differ between the two problems?

Solve the problem.
32) A random sample of 250 students at a university finds that these students take a mean of 15.2 credit hours per quarter with a standard deviation of 2.3 credit hours. Estimate the mean credit hours taken by a student each quarter using a 95% confidence interval. Round to the nearest thousandth.

33) A previous random sample of 4000 U.S. citizens yielded 2250 who are in favor of gun control legislation. How many citizens would need to be sampled for a 90% confidence interval to estimate the true proportion within 1%?

Assume that a hypothesis test of the given claim will be conducted. Identify the type I or type II error for the test.
34) A medical researcher claims that 12% of children suffer from a certain disorder. Identify the type I error for the test.

35) The principal of a school claims that the percentage of students at his school that come from single-parent homes is 16%. Identify the type II error for the test.

Identify the null hypothesis, alternative hypothesis, test statistic, critical value or P-value, conclusion about the null hypothesis, and final conclusion that addresses the original claim.

36) Various temperature measurements are recorded at different times for a particular city. The mean of 20°C is obtained for 60 temperatures on 60 different days. Assuming that $\sigma = 1.5$°C, test the claim that the population mean is 22°C. Use a 0.05 significance level. Construct and Interpret a 95% Confidence interval.

Assume that a simple random sample has been selected from a normally distributed population and test the given claim. Identify the null and alternative hypotheses, test statistic, critical value, and state the final conclusion that addresses the original claim.

37) A large software company gives job applicants a test of programming ability and the mean for that test has been 160 in the past. Twenty-five job applicants are randomly selected from one large university and they produce a mean score and standard deviation of 183 and 12, respectively. Use a 0.05 level of significance to test the claim that this sample comes from a population with a mean score greater than 160. Construct and interpret a 90% cofidence interval.

Test the indicated claim about the means of two populations. Assume that the two samples are independent simple random samples selected from normally distributed populations.

38) A researcher was interested in comparing the resting pulse rates of people who exercise regularly and of those who do not exercise regularly. Independent simple random samples of 16 people who do not exercise regularly and 12 people who exercise regularly were selected, and the resting pulse rates (in beats per minute) were recorded. The summary statistics are as follows.

Do not exercise regularly	Exercise regularly
$\overline{x_1} = 73.1$ beats/min	$\overline{x_2} = 69.2$ beats/min
$s_1 = 10.9$ beats/min	$s_2 = 8.2$ beats/min
$n_1 = 16$	$n_2 = 12$

Use a 0.025 significance level to test the claim that the mean resting pulse rate of people who do not exercise regularly is larger than the mean resting pulse rate of people who exercise regularly.

Test the given claim about the means of two populations. Assume that two dependent samples have been randomly selected from normally distributed populations.

39) A test of abstract reasoning is given to a random sample of students before and after they completed a formal logic course. The results are given below. At the 0.05 significance level, test the claim that the mean score is not affected by the course. Construct and interpret a 95% confidence interval for the mean difference between the before and after scores.

Before	74	83	75	88	84	63	93	84	91	77
After	73	77	70	77	74	67	95	83	84	75

Identify the null hypothesis, alternative hypothesis, test statistic, P-value, conclusion about the null hypothesis, and final conclusion that addresses the original claim.

40) A nationwide study of American homeowners revealed that 64% have one or more lawn mowers. A lawn equipment manufacturer, located in Omaha, feels the estimate is too low for households in Omaha. Can the value 0.64 be rejected if a survey of 490 homes in Omaha yields 331 with one or more lawn mowers? Use $\alpha = 0.05$.

1) A population is the complete collection of all elements. A sample is a subset of elements drawn from a population. A parameter is a numerical measurement describing some characteristic of a population. A statistic is a numerical measurement describing some characteristic of a sample. A census is the collection of data from every element in a population; a sample is a subset of a population.

2) Qualitative data can be separated into categories that are distinguished by nonnumeric characteristics. Quantitative data consist of numbers representing counts or measurements. Examples will vary.

3) Continuous numerical data result from infinitely many possible values that can be associated with points on a continuous scale so that there are no gaps or interruptions. Discrete data result from either a finite number of possible values or a countable number of possible values. Examples will vary.

4) Answers will vary. Possible answer: Histograms convey quantitative information about shapes of distributions. Pareto charts convey comparative information about relative standing of categorical data.

5) By knowing the standard deviation for both groups, we would have an idea about how the individual scores for each group varied about 85. The smaller standard deviation would indicate that individual scores were closer to 85 than would a larger standard deviation.

6) A small variance is preferred, since this measure denotes consistency in the lifetime of the tires. Given small variation, buyers would get useful mileage from those tires around 42,000. Large variation would indicate that some buyers could have their tires wear out many miles short of 42,000, whereas others might get good use out of many miles past 42,000.

7) Marla's score of 85% on her statistics exam tells us that Marla knew 85% of the content on that exam. Marla's percentile score of 85 tells us that her score was better than 85% of the scores of examinees on that test.

8) Central air: mean = $66.20; median = $65
Window unit: mean = $71.60; median = $84
Window units appear to be significantly more expensive.

9) For this data set, the range is very misleading. The range depends only on the smallest and largest values and the remainder of the data contributes nothing to the range. In this case, the smallest value is an outlier. Thus even though all the values except one lie between 89 and 91, the range is 31. The standard deviation, while it will also be affected by the outlier, will be less misleading, as it depends on every piece of data.

10) 2.1; unusual

11) A score of 92

12) 6.4 lb

13) C

14) $\dfrac{3}{4}$

15) 0.42

16) 0.544

17) $\dfrac{43}{48}$

18) 0.993

19) 0.081

20) The four requirements are:
    1) The experiment must have a fixed number of trials.
    2) The trials must be independent.
    3) Each trial must have all outcomes classified into two categories.
    4) The probabilities must remain constant for each trial.
    Answers will vary for the experiment.

21) No, X is not a random variable because it does not have numerical values. A random variable is a variable that has a single numerical value, determined by chance for each outcome of a procedure. Let Y represent the number of heads (or the number of tails) when a coin is flipped three times. Then Y is a random variable with the distribution shown below.

Y (number of heads)	Probability
0	1/8
1	3/8
2	3/8
3	1/8

22) $\mu = 1.23$

23) -$0.67

24) 126

25) 0.207

26) 0.965

27) 2.05

28) $\dfrac{1}{3}$

29) 0.2177

30) 99.12°F

31) In the first, the student must use the formula $z = \dfrac{x - \mu}{\sigma}$, whereas in the second problem, the student must use the formula $z = \dfrac{\overline{x} - \mu}{\dfrac{\sigma}{\sqrt{n}}}$, since the problem involves a sample.

32) $15.2 \pm .285$

33) 6660

34) Reject the claim that the percentage of children who suffer from the disorder is equal to 12% when that percentage is actually 12%.

35) Fail to reject the claim that the percentage of students that come from single-parent homes is equal to 16% when that percentage is actually different from 16%.

36) $H_0: \mu = 22$; $H_1: \mu \neq 22$. Test statistic: $z = -10.33$. P-value: 0.0002. Because the P-value is less than the significance level of $\alpha = 0.05$, we reject the null hypothesis. There is sufficient evidence to warrant rejection of the claim that the population mean temperature is 22°C.

37) $H_0: \mu = 160$. $H_1: \mu > 160$. Test statistic: $t = 9.583$. P-value < 0.005. Reject H_0. There is sufficient evidence to support the claim that the mean is greater than 160.

38) $H_0: \mu_1 = \mu_2$

$H_1: \mu_1 > \mu_2$

Test statistic: $t = 1.080$

Critical value: $t = 2.060$

Do not reject H_0. At the 2.5% significance level, there is not sufficient evidence to support the claim that the mean resting pulse rate of people who do not exercise regularly is larger than the mean resting pulse rate of people who exercise regularly.

39) $H_0: \mu_d = 0$. $H_1: \mu_d \neq 0$.

Test statistic t = 2.366. Critical values: t = ± 2.262.

Reject H_0. There is sufficient evidence to warrant rejection of the claim that the mean is not affected by the course.

40) $H_0: p = 0.64$. $H_1: p > 0.64$. Test statistic: z = 1.64. P-value: p = 0.0505.

Critical value: z = 1.645. Fail to reject null hypothesis. There is not sufficient evidence to warrant rejection of the claim that the proportion with lawn mowers in Omaha is 0.64.