Practice Problems Section 4A

(#1-10) Use each of the following two-population mean T-test statistics and the corresponding critical values to fill out the table.

	Type of Test	T-test stat	Sentence to explain T-test statistic.	Critical Value	Does the T-test statistic fall in a tail determined by a critical value? (Yes or No)	Are the sample means from the two groups significantly different or not? Explain.	Does sample data significantly disagree with H_0? Explain.
1.	Right Tailed	+1.383		+2.447			
2.	Left Tailed	−2.851		−1.773			
3.	Two Tailed	−1.501		±2.006			
4.	Right Tailed	+3.561		+1.692			
5.	Two Tailed	+0.887		±1.943			
6.	Left Tailed	−1.003		−2.759			
7.	Two Tailed	−4.416		±1.994			
8.	Right Tailed	+0.275		+1.839			
9.	Left Tailed	−1.461		−1.674			
10.	Two Tailed	+2.330		±2.138			

(#11-20) Use each of the following P-values and corresponding significance levels to fill out the table.

	P-value Proportion	P-value %	Sentence to explain the P-value	Significance Level %	Significance level Proportion	If H_0 is true, could the sample data occur by sampling variability or is it unlikely?	Reject H_0 or Fail to reject H_0?
11.	0.0007			10%			
12.	0.421			1%			
13.	8.71×10^{-5}			5%			
14.	0.339			1%			
15.	0.076			5%			
16.	0			10%			
17.	0.528			5%			
18.	0.0277			10%			
19.	3.04×10^{-6}			1%			
20.	0.178			5%			

21. Explain the difference between matched pair data and independent groups.

22. Explain the difference between random samples and random assignment.

23. List the assumptions that we need to check for a two-population mean hypothesis test from independent groups.

24. List the assumptions that we need to check for a two-population mean hypothesis test from matched pairs.

25. List the assumptions that we need to check for a two-population mean hypothesis test that is using experimental design.

26. Explain how to use a two-population mean hypothesis test to show that categorical and quantitative data are related.

27. Explain how to use a two-population mean hypothesis test to show there is a cause and effect between categorical and quantitative data.

(#28-33) Directions:

a) Determine if the following two-population mean tests are matched pair or independent groups

b) Write the null and alternative hypothesis. Include relationship implications.

c) Check all of the assumptions for a two-population mean T-test. Explain your answers. Does the problem meets all the assumptions?

d) Write a sentence to explain the T-test statistic.

e) Use the test statistics and the critical value to determine if the sample data significantly disagrees with the null hypothesis. Explain your answer.

f) Write a sentence to explain the P-value.

g) Use the P-value and significance level to determine if the sample data could have occurred by random chance (sampling variability) or is it unlikely to random chance? Explain your answer.

h) Should we reject the null hypothesis or fail to reject the null hypothesis? Explain your answer.

i) Write a conclusion for the hypothesis test. Explain your conclusion in plain language.

j) Is the categorical variable related to the quantitative variable? Explain your answer.

28. The ACT exam is used by many colleges to test the readiness of high school students for college. Many high school students are now taking ACT prep classes. A local high school offers an ACT prep class, but wants to know if it really helps. Twenty students were randomly selected. They took the ACT exam before and after taking the ACT prep class. For each student the difference between the after and before scores were measured (d = after – before). Population 1 was the after prep class scores and population 2 was the before prep class scores. The mean of the differences was 1.5 ACT points with a standard deviation of 2.3 ACT points. A histogram of the differences yielded a bell shaped normal distribution. Use a 5% significance level to test the claim that the after prep class scores are higher than the before prep class scores.   What does this data indicate about the relationship between taking a prep class or not and ACT scores.

N	Sample Mean	Stdev	Significance Level	Critical Value	Test Statistic t	p-Value
20	1.5	2.3	0.05	1.729	2.917	0.0044

29. A random sample of 20 male German Shepherds found that their average weight was 112 pounds with a standard deviation of 28 pounds. A random sample of 14 male Dobermans found that their average weight is 107 pounds with a standard deviation of 24 pounds. Assume that weights are normally distributed. Use the Statcato printout below and a 5% significance level to test the claim that the population mean average weight of male German Shepherds (population 1) is more than the population mean average weight of male Doberman Pinchers (population 2). What does this data indicate about the relationship between the weight and the type of dog?
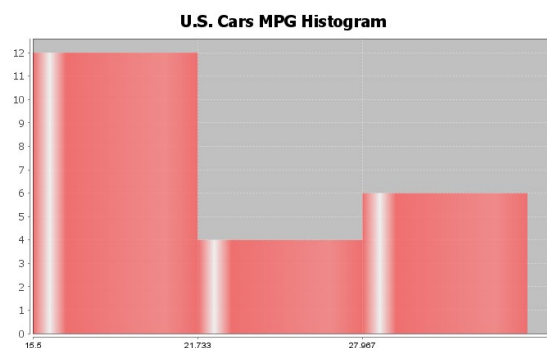
	N	Mean	Stdev
German Shep Sample 1	20	112.0	28.0
Doberman Sample 2	14	107.0	24.0

Significance Level	Critical Value	Test Statistic t	p-Value
0.05	1.697	0.558	0.2906

30. Cotinine is an alkaloid found in tobacco and is used as a biomarker for exposure to cigarette smoke. It is especially useful in examining a person's exposure to second hand smoke. A random sample of 32 non-smoking American adults was collected. These adults were not smokers and did not live with any smokers. The average cotinine level for this sample was 7.2 ng/mL with a standard deviation of 5.8 ng/mL.   A second random sample of 35 non-smoking American adults was then collected. These adults did not smoke themselves, but did live with one or more smokers. The mean average cotinine level for this sample was 28.5 ng/mL and had a standard deviation of 11.4 ng/mL. Use a 1% significance level to test the claim that people that do not live with smokers have a lower cotinine level than those people that do live with smokers. What does this data indicate about the relationship between cotinine levels and living with a smoker or not.

Significance Level	Critical Value	Test Statistic t	p-Value
0.01	-2.402	-9.758	$1.4751 \cdot 10^{-13}$

31. We want to see if the country a car is made in is related to its gas mileage in miles per gallon. Specifically we wanted to see if cars made in the U.S. have a lower population mean average mpg than those made outside the U.S. We used the random car data at www.matt-teachout.org and a 5% significance level to create the following graphs and statistics with Statcato. Check the assumptions and perform the hypothesis test.

U.S. Cars MPG Histogram

Other Country Cars MPG Histogram
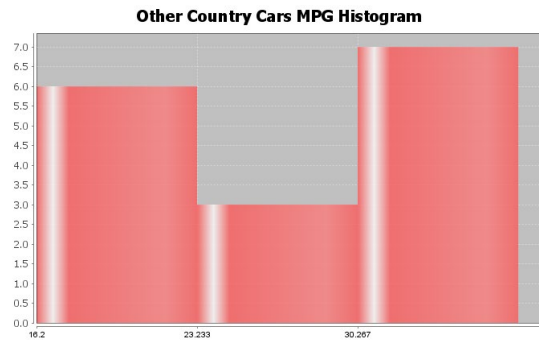
	N	Mean	Stdev
Population 1 USA mpg	22	22.995	6.054
Population 2 Other Country mpg	16	27.188	6.601

* Population standard deviations are unknown. *
DOF = 30

Significance Level	Critical Value	Test Statistic t	p-Value
0.05	-1.697	-2.001	0.0273

32. We want to test the claim that the diastolic blood pressure of a person is less than the systolic blood pressure of a person. We used the random health data at www.matt-teachout.org, Statcato, and a 1% significance level to create the following graphs and statistics. Check the assumptions and perform the hypothesis test. Notice that since the diastolic and systolic blood pressures came from the same randomly selected adults, we used a matched pair calculation.

Diastolic BP Histogram

Systolic BP Histogram
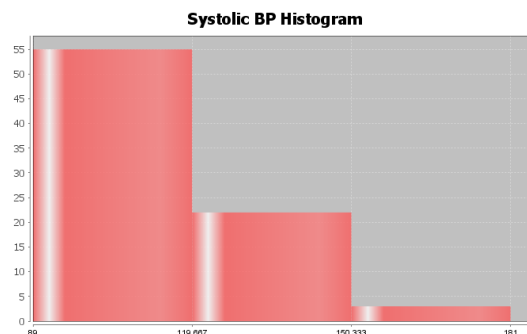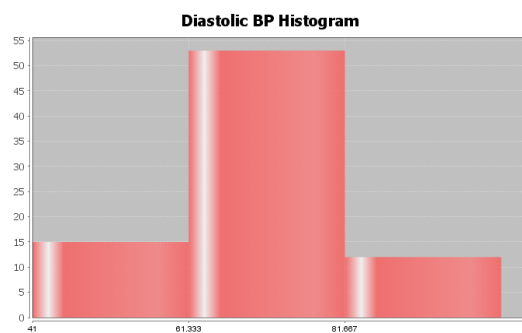
N	Sample Mean	Stdev	Significance Level	Critical Value	Test Statistic t	p-Value
80	-44.525	10.077	0.01	-2.375	-39.521	$4.188 \cdot 10^{-54}$

33. We want to see if the country a car is made in is related to the horsepower of the car. Specifically we wanted to see if cars made in the U.S. have a higher population mean average horsepower than those made outside the U.S. We used the random car data at www.matt-teachout.org and a 10% significance level to create the following graphs and statistics with Statcato. Check the assumptions and perform the hypothesis test.

USA Cars Horsepower Histogram

Other Country Cars Horsepower Histogram
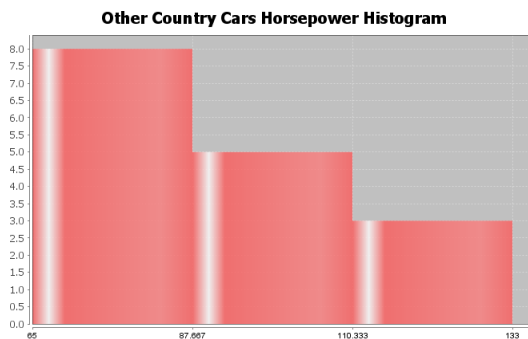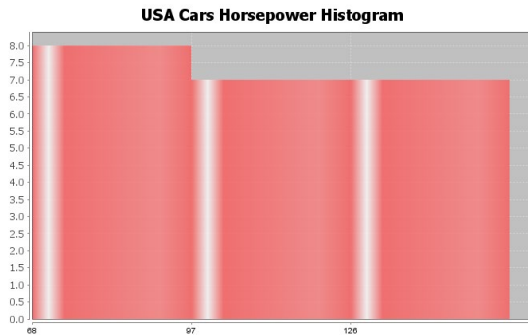
	N	Mean	Stdev
Population 1 USA car horsepower	22	110.182	26.383
Population 2 Other Country car horsepower	16	90.125	22.408

Significance Level	Critical Value	Test Statistic t	p-Value
0.10	1.306	2.526	0.0081

(#34-37) Directions: For each problem, answer the following questions.

a) Determine if the following two-population mean tests are matched pair or independent groups

b) Write the null and alternative hypothesis. Include relationship implications.

c) Use randomized simulation to calculate the P-value. Write a sentence to explain the P-value.

d) Use the P-value and significance level to determine if the sample data could have occurred by random chance (sampling variability) or is it unlikely to random chance? Explain your answer.

e) Use the sample mean difference and the standard error in the simulation to calculate the T-test statistic.

$$T = \frac{Sample\ Mean\ Difference}{Standard\ Error}$$

e) Should we reject the null hypothesis or fail to reject the null hypothesis? Explain your answer.

f) Write a conclusion for the hypothesis test. Explain your conclusion in plain language.

g) Is the categorical variable related to the quantitative variable? Explain your answer.

34. Go to StatKey at www.lock5stat.com. Under the "Randomization Hypothesis Tests" menu, click on "Test for Difference for Difference in Means". Under the data sets menu on the top left, click on "Commute Time (Atlanta vs St Louis)". This took a random sample of people from Atlanta (population 1) and a random sample of people from St. Louis (population 2). Use randomized simulation and a 5% significance level to test the claim that the mean average commute time for people in Atlanta is greater than the mean average commute time for people from St. Louis. What does this data indicate about the relationship between the city and the commute time?

35. Go to StatKey at www.lock5stat.com. Under the "Randomization Hypothesis Tests" menu, click on "Test for Single Mean". Under the data sets menu on the top left, click on "Pulse Rate Difference (Quiz – Lecture)". An experiment was done on college students to determine if heart rate is related to taking a quiz or not. The heart rates of students were measured on a day they were taking a quiz (population 1) and again on a day when there was just lecture (population 2). The same students were measured twice. Use randomized simulation and a 1% significance level to test the claim that the mean average heart rate difference between the quiz and lecture days is greater than zero. This will indicate that the heart rates on quiz days tend to be higher than lecture days. What does this data indicate about the relationship between the heart rate and taking a quiz or not?

36. Go to StatKey at www.lock5stat.com. Under the "Randomization Hypothesis Tests" menu, click on "Test for Difference in Means". Under the data sets menu on the top left, click on "Exercise Hours (Male vs Female)". This took a random sample of male adults (population 1) and a random sample of female adults (population 2). Use randomized simulation and a 10% significance level to test the claim that the mean average amount of time that males and females exercise is the same. What does this data indicate about the relationship between exercise hours and gender?

37. Use StatKey and the random health data to test the claim that the population mean average pulse rate for women is higher than for men. Go to www.matt-teachout.org and click on the statistics tab and then the data sets tab. Open the health data. Copy and paste the gender data and pulse data columns next to each other in a fresh excel spreadsheet. Now copy the two columns. Go to StatKey at www.lock5stat.com. Under the "Randomization Hypothesis Tests" menu, click on "Test for Difference for Difference in Means". Under "Edit Data", paste the gender and pulse rate columns into StatKey. Click on "Generate 1000 Samples" a few times. Click on "Right Tail" and put in the sample mean difference of 6.9 beats per minute in the bottom box in order to estimate the P-value. Now answer the questions above.