Minitab

ACTIVITY SHEET: TWO-SAMPLE T-TESTS FOR POPULATION MEANS
This activity sheet includes exercises to assess students' understanding of important concepts in the *Two-Sample t-Test for Population Means* lesson.

Two-Sample *t*-Test for Population Means

The data for these exercises is in the Minitab file ***TwoSampleTTest_PopMeans_Activity.mtw***.

Exercise 1: Noise Level of Planes

In an environmental impact study for a new airport, the noise level of various jets was measured just seconds after their wheels left the ground. The jets were either wide-bodied or narrow-bodied. The noise levels in decibels (dB) are recorded here for 15 wide-bodied and 12 narrow-bodied jets.*

Wide	109.5	107.3	105.0	117.3	105.4	113.7	121.7	109.2	108.1	106.4	104.6	110.5	110.9	111.0	112.4
Narrow	131.4	126.8	114.1	126.9	108.2	122.0	106.9	116.3	115.5	111.6	124.5	116.2			

* An Introduction to Statistical Methods and Data Analysis
By R. Lyman Ott, Micheal T. Longnecker

(a) Is the data for the wide-bodied and narrow-bodied jets normally distributed so that we can perform a 2-sample *t*-test comparing their noise levels? How are you checking this condition?

(b) Do the two types of jets have different mean noise levels? Set up the null and alternative hypotheses to exam this question. Clearly define the population parameters.

(c) Determine the standardized test statistic and the *p*-value associated with this test statistic.

(d) State the conclusion with respect to the research question.

Exercise 2: Amount Spent on Gifts by Gender

Independent random samples of 23 females and 23 males (no connections to each other) were asked how much money they spent on Valentine's Day gifts this year. Assume the amount spent for each gender is normally distributed and the population variances for the amounts spent are unknown and not equal. A hypothesis test was performed to determine whether the true mean amount that females spent on gifts is equal to the true mean amount that males spent on gifts. The following Minitab output was obtained, but the *p*-value and degrees of freedom did not print.

Sample	N	Mean	StDev	SE Mean
Females	23	25.00	2.00	0.42
Males	23	22.00	3.00	0.63

Test

Null hypothesis	H_0: $\mu_F - \mu_M = 0$
Alternative hypothesis	H_1: $\mu_F - \mu_M \neq 0$

T-Value	DF	P-Value
3.99	xx	x.xxx

What are the degrees of freedom for this hypothesis test, and what is your decision regarding the H_0?

A. We don't need to compute the degrees of freedom because we would use a 2-sample *z*-test to perform this hypothesis test. In performing a 2-sample *z*-test, our decision would be to reject H_0.
B. Without the actual data, we can't determine the degrees of freedom or make a decision about H_0.
C. The degrees of freedom is 22. We would reject H_0.
D. The degrees of freedom is 23. We would reject H_0.
E. The degrees of freedom is 22. We would not reject H_0.
F. The degrees of freedom is 38. We would reject H_0.
G. The degrees of freedom is 44. We would reject H_0.
H. The degrees of freedom is 46. We would not reject H_0.

Exercise 3: Linus's Twitter Followers

Linus is seldom seen without his trusty blue blanket. But he is concerned that it is affecting his social life. Linus decides to conduct a study to determine if children who carry blankets daily have less Twitter followers, on average, than children who do not carry a blanket daily. He obtains a random sample of 20 children who carry a blanket daily and records the number of Twitter followers each has; then, he obtains a random sample of 20 children who do not carry a blanket daily and records the number of Twitter followers each has. The data is provided in Minitab worksheet for this activity document. Note: The order in which the data was collected is unknown.

(a) State the null and alternative hypothesis that addresses the question of interest. Be sure to define the parameters of interest.

(b) Indicate (by circling the choice below) whether you feel the data is paired or independent. Then, compute the corresponding test statistic for assessing the hypothesis stated in part (a).

Paired Data   Two Independent Samples

(c) Woodstock (a big fan of Twitter's mascot) looks over the data Linus has collected and states that "the test statistic in part (b) can be modeled using a standard normal distribution since we have a total of 40 observations randomly collected." Why is this statement invalid?

(d) Assuming all assumptions for modeling the test statistic in part (b) using a *t* distribution are reasonable, Lucy computes a *p*-value that is less than her significance level of $\alpha = 0.05$; she concludes that "his data suggests that if children stopped carrying blankets, then their number of Twitter followers would increase, on average." Is her conclusion justified? Explain.

Exercise 4: Height Data

In class, we examined data from a study conducted in class to estimate the average increase in a student's height when wearing shoes (compared to being barefoot). The data is available in the Minitab worksheet for this activity document.

When designing the study, there was some concern that since males and females wear different types of shoes, the increased height when wearing shoes may differ between genders. As a result, in addition to each student's height with and without shoes, we also recorded each student's gender. We are interested in determining whether the increased height when wearing shoes (compared to being barefoot) differs, on average, between males and females.
Let μ_F be the average increase in height when wearing shoes for females. And, let μ_M be the average increase in height when wearing shoes for males. Then, we are interested in testing the following hypothesis:

$$H_0: \mu_F = \mu_M \text{ versus } H_1: \mu_F \neq \mu_M \text{, or}$$

$$H_0: \mu_F - \mu_M = 0 \text{ versus } H_1: \mu_F - \mu_M \neq 0$$

At first, addressing this question can feel a bit daunting. Since each subject in the study had their height measured with and without shoes, the data is paired (in a sense). This is how we examined the data in class. However, with regard to the question of interest in this problem (does the increased height differ between males and females), the data (that is, the difference in height between wearing shoes and not wearing shoes) is not paired since each male is not paired with a specific female. Again, the idea is that once differences have been taken, the pairing disappears.

Compute the increase in height when wearing shoes for each subject. Tip: Use **Calc > Calculator** to compute the differences. This will be the response moving forward.

Explain why a 2-sample *t*-test is not appropriate for addressing the question of interest. Be sure to include any supporting material (graphics) needed.

Exercise 5: Baby Walkers

Baby walkers are seats hanging from frames that allow babies to sit upright with their legs dangling and feet touching the floor. Walkers have wheels on their legs that allow the infant to propel the walker around the house long before he or she can walk or even crawl. Typically, babies use walkers between the ages of 4 months and 11 months.

Because most walkers have tray tables in front that block babies' views of their feet, child psychologists have begun to question whether walkers affect infants' cognitive development. One study compared mental skills of a random sample of those who never used walkers with a random sample of those who used walkers. Mental skill scores averaged 123 for 55 babies who did not use walkers (standard deviation of 15) and 113 for 54 babies who used walkers (standard deviation of 12).

If we want to construct a 99% two-sided confidence interval for the difference between the true mean mental score for babies who did not use walkers and the true mean mental score for babies who did use walkers, which test would we use?

A. 1-sample *t*-test on differences (or paired *t*)
B. 2-sample *z*-test
C. 2-sample *t*-test

Exercise 6: Ice Cream Comparisons

Suppose you want to determine if, on average, there is more or less fat in one brand of ice cream versus another. So, you take a random sample of 8 tubs of Brand A ice cream and determine the following percentages of fat in these 8 tubs.

A: 5.7, 4.5, 6.2, 6.3, 7.3, 6.1, 5.6, 4.7.

Next, you take a random sample of 8 tubs of Brand B ice cream and determine the following percentages of fat in these 8 tubs.

B: 6.3, 5.7, 5.9, 6.4, 5.1, 5.8, 5.6, 5.7.

Which of the following procedures is most appropriate to test equal versus unequal average fat content in the two types of ice cream?

A. 2-sample *t*-test with 8 degrees of freedom
B. Paired *t* test or 1-sample *t* test on differences with 7 degrees of freedom
C. 2-sample *t*-test with 9 degrees of freedom
D. Two sample *z*-test
E. 2-sample *t*-test with 16 degrees of freedom
F. None of these methods is correct

Exercise 7: Weight-Loss Results

To test the effectiveness of the new weight-loss drug "Reducts," 40 women were randomly split into two groups (e.g. names drawn out of a hat): Group A, the control group, took a placebo drug, and Group B, the experimental group, took Reducts. The amount of weight lost by each member of the groups over a 6-month period is given below.

Group A: 3, 4, 5, 6, 7, 10, 11, 12, 15, 18, 19, 20, 23, 24, 25, 30, 33, 38, 40, 42
Group B: 1, 2, 5, 7, 9, 10, 12, 14, 15, 18, 20, 24, 28, 29, 38, 42, 44, 49, 50, 55

Based on these results, is Reducts ***effective*** (i.e. increased weight loss)? We'll run a hypothesis test to determine this.

Using the appropriate hypothesis test, provide the standardized test statistic and corresponding *p*-value for that test statistic. Make sure that all necessary assumptions have been met to run the given hypothesis test that you have chosen.

Exercise 8: Lab vs Golden Retriever Weights

After seeing a picture of my student's Labrador Retriever on his computer desktop, I was curious as to whether Labrador Retrievers (L) or Golden Retrievers (G) weigh more as adult dogs. So, I decided to collect data in order to determine if their true mean weights are the same or not. The data I collected for each breed is in the Minitab worksheet for this activity. Perform the following hypothesis test.

$$H_0: \mu_L = \mu_G \text{ versus } H_a: \mu_L \neq \mu_G$$

(a) Is the data paired or independent?

(b) Determine the standardized test statistic.

(c) Determine the *p*-value associated with this test statistic.

(d) At $\alpha = 0.05$, can we reject H_0?

Exercise 9: Mice Maze Times

The times (in minutes) it took six white mice to complete a maze and the times it took six brown mice to complete the same maze are given below. At $\alpha = 0.05$, does the color of the mice make a difference in their learning rate?

White Mice	18	24	20	13	15	12
Brown Mice	25	16	19	14	16	10

Use a 95% confidence interval for the difference of their mean times to complete the maze to answer this question. Assume the times for both groups are from normally distributed populations.

Exercise 10: Employee Sick Days

A large corporation is interested in determining whether the average days of sick leave taken annually is more for the nightshift employees than for the dayshift employees. It is assumed that the distribution of the days of sick leave is normal for both shifts and that the variances are unknown and assumed unequal.

A random sample of 12 employees from the nightshift yields an average sick leave of 16.4 days with a standard deviation of 2.2 days. A random sample of 15 employees from the dayshift yields an average sick leave of 12.3 days with a standard deviation of 3.5 days.

The hypotheses are: H_0: $\mu_{NS} = \mu_{DS}$ vs H_1: $\mu_{NS} > \mu_{DS}$, or H_0: $\mu_{NS} - \mu_{DS} = 0$ vs H_1: $\mu_{NS} - \mu_{DS} > 0$.

(a) Determine the standardized test statistic for the hypothesis test.

(b) Should we reject H_0 at level of significance $\alpha = 0.01$?

Exercise 11: DDT Found in Pelicans

A study is being conducted by the Florida Game and Fish Commission to assess the amounts of chemical residues found in the brain tissue of brown pelicans. Specifically, the commission is interested in determining if the mean amount of DDT found in juvenile pelicans is larger than the mean amount of DDT found in nestling pelicans. [A nestling pelican is one that is between the hatching and leaving the nest period.] This test has important implications regarding the accumulation of DDT over time.

A random sample of 10 juvenile pelicans is selected and the DDT amounts in parts per million (ppm) are recorded in the column "Juvenile" in the Minitab worksheet for this activity. Likewise, a random sample of 10 nestling pelicans is selected and the DDT amounts in parts per million (ppm) are in the column "Nestling."

Let $\mu_{juvenile}$ be the true average amount of DDT (in ppm) found in juvenile pelicans in this area, $\mu_{nestling}$ be the true average amount of DDT (in ppm) found in nestling pelicans in this area, and μ_{diff} be the true mean difference in the amount of DDT in juveniles versus nestlings. Specifically, let "diff = juvenile – nestling."

(a) Select the appropriate null and alternative hypotheses for this experiment using the parameters defined in the above paragraph. Since the alternative can be written in various ways depending on the order of subtraction for the variable "diff," there may be more than one correct answer. Select all the correct choices.

A. H_0: $\mu_{juvenile} - \mu_{nestling} = 0$ versus H_a: $\mu_{juvenile} - \mu_{nestling} < 0$
B. H_0: $\mu_{nestling} - \mu_{juvenile} = 0$ versus H_a: $\mu_{nestling} - \mu_{juvenile} > 0$
C. H_0: $\mu_{nestling} - \mu_{juvenile} = 0$ versus H_a: $\mu_{nestling} - \mu_{juvenile} \neq 0$
D. H_0: $\mu_{juvenile} - \mu_{nestling} = 0$ versus H_a: $\mu_{juvenile} - \mu_{nestling} > 0$
E. H_0: $\mu_{diff} = 0$ versus H_a: $\mu_{diff} < 0$
F. H_0: $\mu_{juvenile} - \mu_{nestling} > 0$ versus H_a: $\mu_{juvenile} - \mu_{nestling} = 0$

(b) Is the data paired or independent?

(c) Determine the standardized test statistic for the appropriate hypothesis test.

(d) What is the *p*-value for the test you have conducted?

(e) Based on the significance level $\alpha = 0.05$, what decision should the commission reach?

(f) Determine the correct *t* critical value (*cv*) in the expression below to construct a 98% two-tailed confidence interval for the difference in the mean amounts of DDT.

$$0.0025 \pm \boldsymbol{cv} \cdot \sqrt{\frac{0.00465^2}{10} + \frac{0.00254^2}{10}}$$

Exercise 12: Drinking Water and Weight Loss

A new study published in the journal *Obesity* found that pre-loading water *before* meals helps you lose weight. The study looked at 84 obese adults and had 41 members of the group drink around 16 ounces of water before meals, while the other 43 adults were asked simply to imagine being full before digging into their food.

Over the course of the 12 weeks, those who filled up on water prior to eating the three main meals a day lost an average of 9.48 pounds with a standard deviation of 3.12 pounds, whereas those imagining they were full before meals resulted in an average loss of 1.76 pounds with a standard deviation of 0.21 pounds.

Let μ_1 be the average weight loss for the participants drinking 16 ounces of water before meals and let μ_2 be the average weight loss for participants who imagined they were full before meals. The authors were interested in results of the following hypothesis test. Their intention was to show that the average weight loss for participants who drank water prior to meals was greater than the average weight loss for participants imagining they were full before meals.

$$H_0: \mu_1 - \mu_2 \le 0 \qquad \text{vs.} \qquad H_1: \mu_1 - \mu_2 > 0$$

(a) Compute an appropriate confidence interval for addressing this hypothesis at $\alpha = 0.025$ significance level. Write down *both* the formula and the final answer. Use the *most* appropriate method to model the underlying distribution, e.g. normal, Student's *t*, *F*, …

(b) Given the confidence interval that you computed in part (a), is it possible that the difference in mean weight loss amounts could be as much as 8.5 pounds at $\alpha = 0.025$?

Exercise 13: Fish Tank Filters from "Finding Nemo"

After the "Tank Gang's" escape, the dentist P. Sherman decides to compare two filters for use in his fish tank. Specifically, he is interested in determining if there is evidence that a power filter will remove a higher percentage of contaminants compared to an internal filter. He takes a sample of 40 fish tanks and observes the type of filter currently in use and the percentage of contaminants removed. Below is a summary of the data.

Filter	**N**	**Mean**	**Std. Dev.**
Power	20	98.5	6.91
Internal	20	92.3	10.63
Diff. (Power – Internal)	20	6.2	8.27

(a) State the null and alternative hypothesis that best addresses the question of interest. Be sure to define any mathematical notation used.

(b) Compute an appropriate confidence interval for addressing the hypothesis stated in part (a) at the 0.01 significance level. Be sure to state *both* the formula and the final answer. You may assume that the percentage of contaminants removed by each filter is from a normal distribution.

(c) Using the interval constructed in part (b), what conclusions can be drawn? Be sure to state your conclusions in the context of the problem.