Name: ____ Due Monday, April 7th

AP Statistics | Baltimore Polytechnic Institute

Paired t-Tests

Hi guys! I miss you already! Let's learn about paired t-tests and feel really great that we're learning our FIFTH method for inference. I know! I'm excited too! Read carefully and do what you are asked. This will be graded.

We'll jump right into an example, an example similar to the one we looked at yesterday:

One indicator of physical fitness is resting pulse rate. Ten men volunteered to test an exercise device advertised on television by using it three times a week for 20 minutes. Their resting pulse rate (beats per minute) were measured before the test began, and then again after six weeks. Results are shown in the table. Is there evidence that this kind of exercise can reduce resting pulse rates? How much?

	Pulse rates (beats/min)	
Subject	**Before**	**After**
Allen	73	73
Brandon	83	79
Carlos	85	81
David	87	86
Edwin	91	87
Franco	99	91
Graeme	87	84
Hans	85	83
Ivan	83	84
Jorge	79	76

Let's do something crazy and start by checking the conditions first, even before we write our hypotheses. So, go ahead and check just two of the conditions you would check for a 2-sample t-test.

1. Random

2. Nearly Normal (skip this condition for the moment)

3. Independent Groups

1. Though we don't have a random sample, we can assume that these ten guys are representative of dudes in general. You would probably want to repeat the study with a lot more people, but this is OK for now.

2. I told you to skip this condition. You'll see why in a moment. I promise.

3. Dang. These two groups are definitely NOT independent. The "after" pulse rate is related to the "before" pulse rate for each of these guys. And if the groups are not independent, we CANNOT EVEN PROCEED WITH CAUTION. That's right. I said it. You cannot do a 2-sample t-test here because the groups are not independent.

Why is the independence condition so important here? (Hint: think about standard error)

The two groups are not independent, but the data is **paired.** In an experiment, this pairing comes from a type of *blocking*. In an observational study, this pairing is a form of *matching*. Is the example problem an experiment or an observational study? ____

Paired t-test

In a paired t-test, what we're most interested in looking at is the *differences* between the pairs. Go ahead and calculate the difference between the pulse rates for each dude. I've repeated the table below and I've even started it for you! You fill in the rest. Go ahead, I'll wait.

Subject	Before (bpm)	After (bpm)	Difference (Before – After)
Allen	73	73	73 – 73 = 0
Brandon	83	79	83 – 70 = -4
Carlos	85	81	
David	87	86	
Edwin	91	87	
Franco	99	91	
Graeme	87	84	
Hans	85	83	
Ivan	83	84	
Jorge	79	76	

What does a positive difference indicate in this context?

What does a negative difference indicate in this context?

Now go ahead and find the average and standard deviation of the differences. To be fancy, we write the difference average as $\bar{d}$. We write the standard deviation of the differences as s_d, which isn't as fancy. Sorry.

Calculations: $\bar{d} =$ $s_d =$

Hypotheses

The null hypothesis will be: $H_0: \mu_d = \Delta_0$,
where Δ_0 is generally 0, meaning that there is no difference.

The null hypothesis will be: $H_A: \mu_d > or < or \neq \Delta_0$,
depending on the question being asked

Conditions

1. Randomization. But, you probably could have guessed that. Let's look at how randomness can occur:

- The order of two treatments might be randomly assigned in an experiment.
- The treatments may be randomly assigned to one member of each pair.
- In a before/after study, we may believe that the observed differences are a representative sample from a population of interest

Which one of these applies for our example about resting pulse?

2. Nearly Normal. This condition can be checked with a histogram or boxplot of the *differences* - but not of the individual groups. This is why I didn't have you check condition 2 at the very beginning. Go ahead and check it now, making just one boxplot or histogram for the *differences*.

3. Paired Data Assumption: You need to state that the data is paired, and how you know it's is paired.

How do you know the data is paired for our pulse rates example?

Mechanics

The mechanics are easy! They're easy because you're not doing anything different from a 1-sample t-test. You just have one statistic to look at, and that's $\bar{d}$, the mean of the paired differences.

$$t_{n-1} = \frac{\bar{d} - 0}{SE(\bar{d})}$$

where $\bar{d}$ is the mean of the pairwise differences and n is the number of *pairs*.

$$SE(\bar{d}) = \frac{s_d}{\sqrt{n}}$$

The interval for a paired t-test is also really simple

$$\bar{d} \pm t_{n-1}^{*} SE(\bar{d})$$

You'll notice in that in your calculator, there isn't a paired t-test to choose in STAT→TESTS. Instead, you will do a TTEST of the differences, which makes sense since we're really just calculating a 1-sample t-test or interval but for the differences of the pair.

Conclusion

When you write your conclusion, you must discuss the *difference*. Same old stuff here though. Reject or fail to reject, tie your decision to an α level, and state whether there is strong or insufficient evidence, all in context.

Check out the back for practice. Bring these with you to class on Monday!

Now try this:

1. In developed countries, the average age of women is generally higher than that of men. After all, women tend to live longer. But if we look at married couples, husbands tend to be slightly older than wives. How much older, on average, are husbands? Create a 90% confidence interval based on a random sample of 170 couples, whose average difference in age was found to be 2.2 years with a standard deviation of 4.1 years.

2. What are the advantages and disadvantages of increasing our confidence level to 95%? Specifically discuss types of error and power.

3. Having done poorly on their Math final exams in June, six students repeat the course in summer school and take another exam in August. Here are their results:

June	54	49	68	66	62	62
August	50	65	74	64	68	72

a) If we consider these students to be representative of all students who might attend summer school in other years, do these results provide evidence that the program is worthwhile?

b) This conclusion, of course, may be incorrect. If so, which type of error was made?