Lagunita is retiring and will shut down at 12 noon Pacific Time on March 31, 2020. A few courses may be open for self-enrollment for a limited time. We will continue to offer courses on other online learning platforms; visit http://online.stanford.edu.

Matched Pairs: Hypotheses

Learning Objective: Specify the null and alternative hypotheses for comparing groups.

Step 1: Stating the hypotheses.

Recall that in the t-test for a single mean our null hypothesis was: $H_o : \mu = \mu_o$ and the alternative was one of $H_a : \mu < or > or \neq \mu_0$. Since the paired t-test is a special case of the one-sample t-test, the hypotheses are the same except that:

- Instead of simply μ we use the notation μ_d to denote that the parameter of interest is the mean of the differences.
- In this course our null value μ_0 is always 0 (although technically, it does not have to be).

Therefore, in the paired t-test:

The null hypothesis is always:

$$H_o : \mu_d = 0$$

and the alternative is one of :

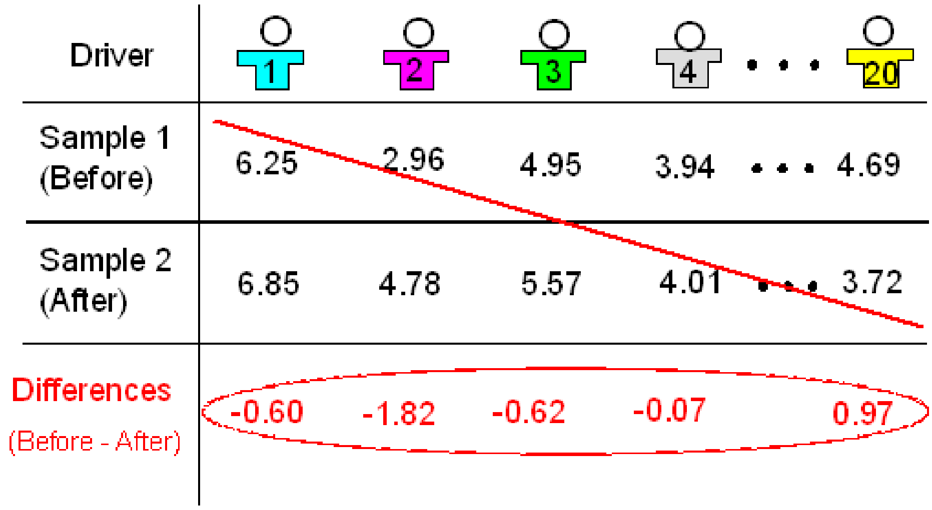
$$\begin{aligned} H_a &: \mu_d < 0 \quad (one-sided) \\ H_a &: \mu_d > 0 \quad (one-sided) \\ H_a &: \mu_d \neq 0 \quad (two-sided) \end{aligned}$$

depending on the context.

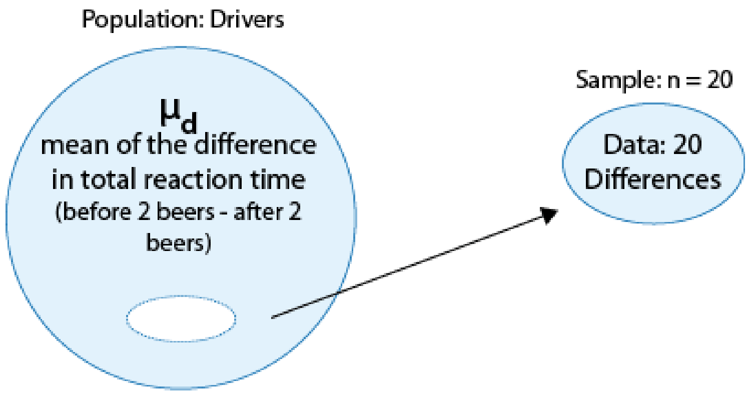
Let's go back to our example to see how this works and why it makes sense.

Example: Drunk Driving

Recall that in our "Are drivers impaired after drinking two beers?" example, our data was reduced to one sample of differences (one for each driver),



so our problem was reduced to inference about the mean of the differences μ_d .



As we mentioned, the null hypothesis is:

$$H_o : \mu_d = 0 .$$

The null hypothesis claims that the differences in reaction times are centered at (or around) 0, indicating that drinking two beers has no real impact on reaction times. In other words, drivers are not impaired after drinking two beers.

In order to decide which of the alternatives is appropriate here we have to think about the context of the problem. Recall that we want to check whether drivers are impaired after drinking two beers. Thus, we want to know whether their reaction times are longer after the two beers. Since the differences were calculated before-after, longer reaction times after the beers would translate into negative differences. These differences are: 6.25 - 6.85, 2.96 - 4.78, etc.

Therefore, the appropriate alternative here is:

$$H_a : \mu_d < 0$$

indicating that the differences are centered at a negative number.

Many Students Wonder ...

Question: Many students wonder whether in matched-pairs analysis you should always calculate the differences "response 1 - response 2," or whether it is possible to calculate and do the analysis using the differences "response 2 - response 1."

Answer: Both ways of calculating the differences are fine (as long, of course, as all the differences are calculated the same way). You should pay careful attention to how the differences are calculated, since that will determine the direction of the alternative hypothesis (in the one-sided case). In our example, the differences in reaction times were calculated: before 2 beers - after 2 beers, and therefore in order to test whether drivers are impaired after two beers (i.e., whether their reaction times are longer after drinking two beers) the appropriate hypotheses are:

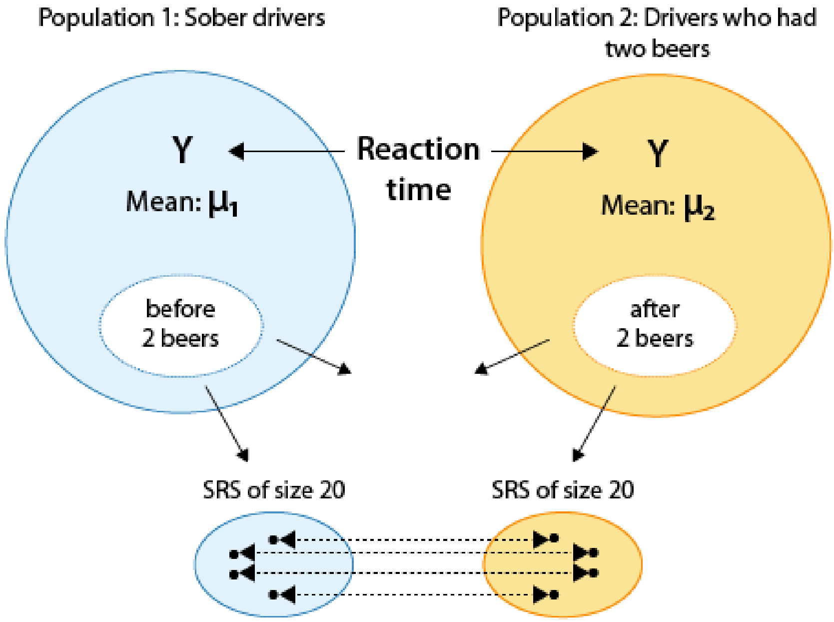
$$H_0 : \mu_d = 0 \quad \text{vs.} \quad H_a : \mu_d < 0$$

However, if the differences were calculated: after 2 beers - before 2 beers, the appropriate alternative hypothesis would also change direction, and we would test:

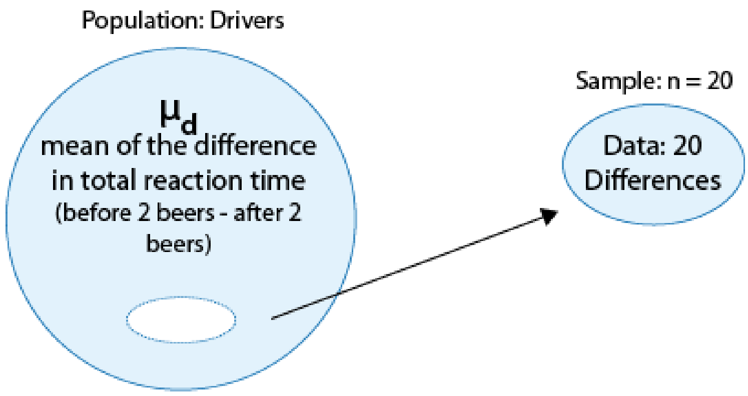
$$H_0 : \mu_d = 0 \quad \text{vs.} \quad H_a : \mu_d > 0$$

Comment

Recall that originally, the following figure represented our problem:



Later, we reduced the problem to inference about a single mean, the mean of the differences:



Some students find it helpful to know that it turns out that $\mu_d = \mu_1 - \mu_2$. In other words, the difference between the means $\mu_1 - \mu_2$ in the first representation is the same as the mean of the differences, μ_d,in the second one. Some students find it easier to first think about the hypotheses in terms of $\mu_1 - \mu_2$ (as we did in the two-sample case) and then represent it in terms of μ_d.

In our example, since we want to test whether the reaction times in population 1 are shorter, we are testing $H_o : \mu_1 - \mu_2 = 0 \ vs. \ H_a : \mu_1 - \mu_2 < 0$, which in the matched pairs design notation is translated to $H_o : \mu_d = 0 \ vs. \ H_a : \mu_d < 0$.

Here is another example:

Example

Suppose the effectiveness of a low-carb diet is studied with a matched pairs design, recording each participant's weight before and after dieting. What would be the appropriate hypotheses in this case?

As before, μ_d is the mean of the differences (weight before diet)-(weight after diet). In this case, if the diet is effective and participants' weight after the diet was indeed lower, we would expect the differences to be positive, and therefore the appropriate hypotheses in this case are:
$H_o : \mu_d = 0 \ vs. \ H_a : \mu_d > 0$.

Did I Get This

1/1 point (graded)
In the following case, decide based on the context what the appropriate set of hypotheses is.

In order to test the effectiveness of a new drug in reducing cholesterol level, a random sample of 45 patients who have a higher than normal cholesterol level was chosen. The cholesterol level of each of the patients was measured and recorded before and then after taking the new drug for a period of 6 weeks, and the differences (before - after) were calculated. The appropriate set of hypotheses, in this case, is:

- (H_o: $\mu_d = 0$) and (H_a: $\mu_d < 0$)
- (H_o: $\mu_d = 0$) and (H_a: $\mu_d > 0$) ✔
- (H_o: $\mu_d = 0$) and (H_a: $\mu_d \neq 0$)
- None of the above. This is not a matched pairs situation.

Answer
Correct:
Indeed, if the drug is effective, we would expect the participants' cholesterol levels to be lower after taking it for 6 weeks. This means that we expect the differences (level before - level after) to be positive, and therefore the appropriate hypotheses in this case are (H_o: μ_d = 0) and (H_a: μ_d > 0).

Submit

Did I Get This

1/1 point (graded)
In the following case, decide based on the context what the appropriate set of hypotheses is.

A publishing company wanted to test whether typing speed differs when using word processor A or word processor B. The typing speeds (in words per minute) are recorded for a random sample of 25 typists using word processor A, and for another (different) random sample of 25 typists using word processor B. The appropriate set of hypotheses in this case is:

- (H_o: $\mu_d = 0$) and (H_a: $\mu_d < 0$)
- (H_o: $\mu_d = 0$) and (H_a: $\mu_d > 0$)
- (H_o: $\mu_d = 0$) and (H_a: $\mu_d \neq 0$)
- None of the above. This is not a matched pairs situation. ✔

Answer
Correct:
Indeed this is not a matched pairs situation. Since two (different) random samples of secretaries were chosen, these two samples are independent, and not matched.

Submit

Did I Get This

1/1 point (graded)
In the following case, decide based on the context what the appropriate set of hypotheses is.

A publishing company wanted to test whether typing speed differs when using word processor A or word processor B. A random sample of 25 typists was selected and the typing speeds (in words per minute) were recorded for each typist when using word processor A and then when using word processor B. (Which word processor is used first is determined for each secretary by a coin flip).

- (H_o: $\mu_d = 0$) and (H_a: $\mu_d < 0$)
- (H_o: $\mu_d = 0$) and (H_a: $\mu_d > 0$)
- (H_o: $\mu_d = 0$) and (H_a: $\mu_d \neq 0$) ✔
- None of the above. This is not a matched pairs situation.

Answer
Correct:
Indeed, since we would like to test whether the typing speeds differ, the appropriate hypotheses are: (H_o: $\mu_d = 0$) and (H_a: $\mu_d \neq 0$).

Submit

Open Learning Initiative

Unless otherwise noted this work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

© All Rights Reserved