MatcheLearn By Doing: Matched Pairs

The purpose of this activity is to give you guided practice in carrying out the **paired t-test**.

Here is some background for the historically important data that we are going to work with in this activity.

Background: Gosset's Seed Plot Data (Seeds: SPSS format, SAS format, Excel format, CSV format)

William S. Gosset was employed by the Guinness brewing company of Dublin. Sample sizes available for experimentation in brewing were necessarily small, and new techniques for handling the resulting data were needed. Gosset consulted Karl Pearson (1857-1936) of University College in London, who told him that the current state of knowledge was unsatisfactory. Gosset undertook a course of study under Pearson and the outcome of his study was perhaps the most famous paper in statistical literature, "The Probable Error of a Mean" (1908), which introduced the
t- distribution.

Since Gosset was contractually bound by Guinness, he published under a pseudonym, "Student"; hence, the t distribution is often referred to as **Student's t distribution**.

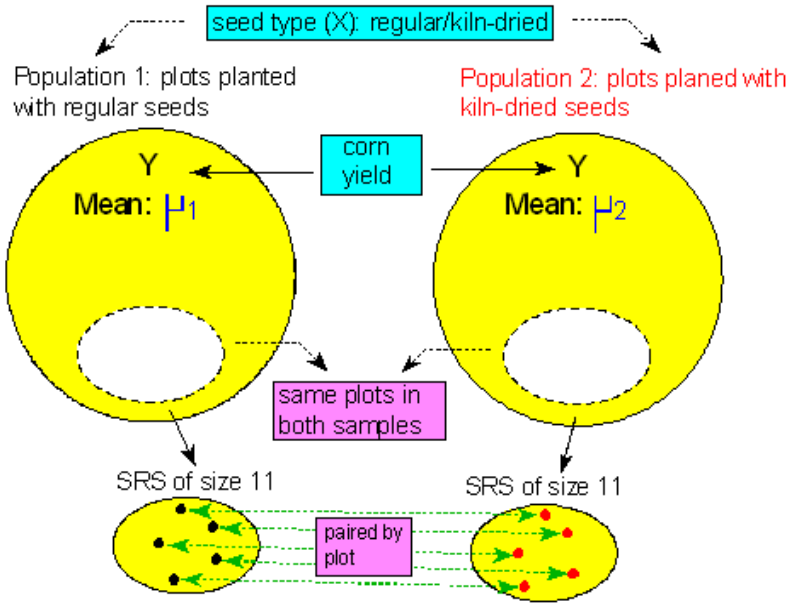
As an example to illustrate his analysis, Gosset reported in his paper on the results of seeding 11 different plots of land with two different types of seed: regular and kiln-dried. There is reason to believe that drying seeds before planting will increase plant yield. Since different plots of soil may be naturally more fertile, this confounding variable was eliminated by using the matched pairs design and planting both types of seed in all 11 plots.

The resulting data (corn yield in pounds per acre) are as follows:

Plot	Regular seed	Kiln-dried seed
1	1903	2009
2	1935	1915
3	1910	2011
4	2496	2463
5	2108	2180
6	1961	1925
7	2060	2122
8	1444	1482
9	1612	1542
10	1316	1443
11	1511	1535

We will use these data to **test the hypothesis that the type of seed is associated with the plant yield**.

Here is a figure that summarizes this problem:

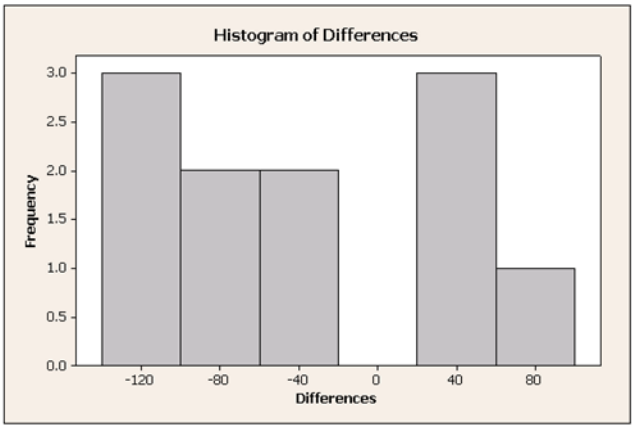


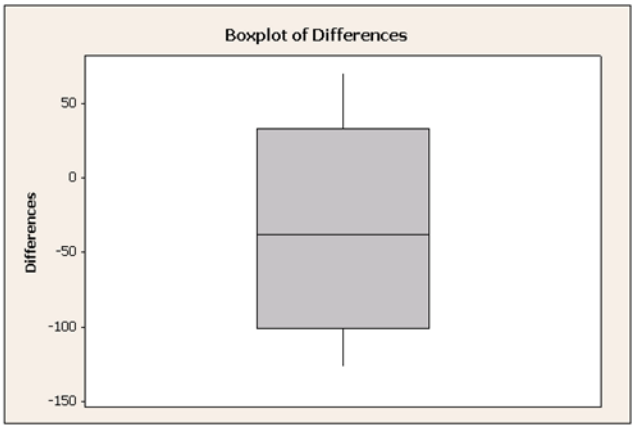
Because of the nature of the experimental design (matched pairs), we are testing the difference in yield.

Plot	Regular seed	Kiln-dried seed	Difference
1	1903	2009	**-106**
2	1935	1915	**20**
3	1910	2011	**-101**
4	2496	2463	**33**
5	2108	2180	**-72**
6	1961	1925	**36**
7	2060	2122	**-62**
8	1444	1482	**-38**
9	1612	1542	**70**
10	1316	1443	**-127**
11	1511	1535	**-24**

Note: the differences were calculated: **regular – kiln-dried**.

Here is a histogram and boxplot of the differences:





To motivate you to conduct the test yourself in your own software, we are providing only this hand-made EXCEL output for this Paired t-test.

t-Test: Paired Two Sample for Means		
	Variable 1	*Variable 2*
Mean	1841.454545	1875.181818
Variance	117468.8727	110789.1636
Observations	11	11
Pearson Correlation	0.981237487	
Hypothesized Mean Difference	0	
df	10	
t Stat	-1.690476076	
P(T<=t) one-tail	0.060908298	
t Critical one-tail	1.812461505	
P(T<=t) two-tail	0.121816596	
t Critical two-tail	2.228139238	

Question 1:

State the appropriate hypotheses that are being tested here. Be sure to define the parameter that you are using.

Your answer

REVEAL OUR ANSWER

Let μ_d be the mean of the yield difference "regular – kiln-dried" in a plot.

As usual, we will use a two-sided test and use the confidence interval to provide more information about the plausible values of the mean difference in yield (regular – kiln-dried) so the appropriate hypotheses are:

H_0: $\mu_d = 0$

H_a: $\mu_d \neq 0$

Question 2:

Are the conditions that allow us to safely use the paired t-test satisfied?

Support your answer by using appropriate visual displays.

Your answer

REVEAL OUR ANSWER

Let's check the conditions:

(i) Even though the problem does not state it specifically, we can assume that Gosset "knew what he was doing" and the plots were chosen at random.

(ii) The sample size is quite small (n = 11), so we need to make sure that the data (differences) do not display any extreme departure from the normality assumption. Typically we look at a histogram.

Comment: some prefer to look at a boxplot of the data, since outliers are clearly marked, and skewness is apparent by the location of the middle line inside the box (which represents the location of the median relative to the quartiles). We recommend looking at the histogram (as we did previously) or at both graphs. See the histogram and boxplot of the data.

Both graphs show us that the data does not display any departure from the normality assumption in the form of extreme skewness and/or outliers.

Histogram: With such a small sample size of n = 11 we cannot really expect to see a normal shape in the histogram; so as long as we don't see anything that is extremely "non-normal," we're fine.

Boxplot: shows us that no observation was classified as an outlier, and that there is no extreme skewness, since the median – the line inside the box – is roughly in the middle of the box).

Question 3:

Based on the visual displays for answering the previous question, does it seem like there is some evidence in the data in favor of the alternative hypothesis? Explain.

Your answer

REVEAL OUR ANSWER

We notice that most of the differences (7 out of 11) are negative, indicating that in 7 of the 11 plots, the dry seeds produced more corn yield. This is evidence in favor of the alternative hypothesis of a difference, but the evidence is not overwhelming, so it is hard to say whether this is strong enough evidence to reject the null hypothesis in favor of the alternative.

Question 4:

Carry out the paired t-test, state the test statistic and p-value, and state your conclusion in context.

Your answer

REVEAL OUR ANSWER

The **test statistic is -1.69 and the p-value is 0.122**, indicating that there is a 12.2% chance of obtaining data like those observed (or even more extremely in favor of the alternative hypothesis – i.e. in either direction) had there really been no difference between regular and kiln-dried seeds (as the null hypothesis claims).

The p-value is larger than our significance level (cut-off probability) of 0.05.

Conclusion: The data do not provide enough evidence that yield differs between kiln-dried seeds and regular seeds.

Question 5:

The 95% confidence interval for the population mean difference (regular – kiln-dried) is:

(-78.1816, 10.7271)

Interpret this interval in terms of the two population means.

Your answer

REVEAL OUR ANSWER

We are 95% confident that the population mean corn yield for kiln-dried seed is between 78.2 pounds per acre greater than that for regular seed to 10.73 pounds per acre less than that for regular seed.

OR

We are 95% confident that the population mean corn yield for regular seed is between 78.2 pounds per acre less than that for kiln-dried seed to 10.73 pounds per acre greater than that for kiln-dried seed.

REVEAL OUR FINAL COMMENTS:

We were not able to reject the null hypothesis – in other words, the difference observed was not statistically significant.

However, the confidence interval does "lean" towards an indication that the kiln-dried seed may increase yield, particularly given the very small sample size of 11 plots.

Additional experiments might confirm this or they may confirm our current result where we found no statistically significant difference.

The mean difference is -33.7 indicating that, in this study, the kiln-dried seeds did **increase corn yield** by 33.7 pounds per acre. We would need to carefully consider whether this is a **practically significant** increase which requires further study or now.