Testing for Differences between two brands of Cookies

KEYWORDS:
Teaching;
2-Sample t-test;
Paired t-test.

Rhonda C. Magel
North Dakota State University,
Fargo, USA.
e-mail: magel@prairie.nodak.edu

Summary
This article presents two class activities for use with hypothesis testing. The first has students collecting data to test for a difference in the mean number of chocolate chips per cookie between two brands. Data are also collected to test for a difference in taste ratings.

◆INTRODUCTION ◆

THE DEPARTMENT of Statistics at North Dakota State University offers several sections of an introductory statistics class every semester. The prerequisite for the class is a knowledge of algebra. It is not assumed that the students have any prior knowledge of statistics. The purpose of the class is to introduce students to the area of statistics and its applications to real-life situations.

Students taking this class are primarily in their second or third year. Their major areas vary including biology, agriculture, food and nutrition, aeronautical engineering, psychology, sociology, business, and child development. Motivating the students requires a great deal of work.

I teach at least one section of this introductory statistics class every year. When introducing various hypothesis tests and confidence intervals in my class, I like to have some activities that require students to collect data, perform a statistical test, perhaps calculate the corresponding confidence interval, and interpret the results. I find that this is a great time to get students actively involved in the classroom. This article discusses two class activities in which data were collected for conducting statistical tests.

◆CLASS ACTIVITY ONE◆

Last semester, after introducing the independent two sample *t*-test, paired *t*-test, corresponding confidence intervals, and assumptions necessary for conducting the tests, I decided to have my students do two activities involving chocolate chip cookies. For the first activity, students in the class were asked to gather data in order to test whether there was evidence to indicate the mean number of chocolate chips per cookie in each of two different brands of cookies was different. I had purchased packages of two kinds of chocolate chip cookies ahead of time, each costing about the same price. Package labels were ripped off the cookies before students saw them. The cookies were only marked as Brand A and Brand B. Every student (there were 22) was given a cookie of each kind and asked to count the number of chips in each. The cookies were of approximately the same size. No formal weight was taken by which to know which cookie actually weighed more or whether their weights were approximately the same. We were just testing whether the mean number of chocolate chips per cookie as made was the same, or not.

The problem of counting the number of chocolate chips in each cookie was discussed, and techniques were mentioned as to how this could be done. For our purposes however, we ended up counting the chips by breaking the dry cookies apart with our hands since we did not have access to a lab. Students could work together in doing the counting. This was actually encouraged to ensure greater accuracy. I mentioned to the class that I had everyone counting chips to get them involved, but the count should not depend on the person doing it.

Students were given the following sheet on which to record their results (Example 1).

Example 1.

Let μ_A = mean number of chocolate chips in each Brand A cookie.
Let μ_B = mean number of chocolate chips in each Brand B cookie.

We would like to test (at $\alpha = 0.05$) to see if there is evidence that the two means differ. A random sample of Brand A cookies was taken and the number of chocolate chips in each cookie was counted with the following results:

Similarly, a random sample of Brand B cookies gave the following results:

NOTE: The class obtained the following data:

Sample A:
21, 14, 23, 15,7, 16, 15, 15, 21, 13, 12, 15, 12, 9, 19, 7, 11, 21, 16, 16, 15, 12.
Sample B:
15, 21, 22, 24, 20, 14, 12, 21, 20, 9, 20, 17, 30, 23, 19, 20, 20, 20, 22, 18, 22, 25.

Later, students took the data to the computer laboratory and completed Lab Sheet 1. Minitab was used to do the calculations.

The first part of Lab Sheet 1 has students checking the assumptions of the independent two-sample *t*-test. Besides having two independent random samples of chip counts, the chip counts from the Brand A cookies should be approximately Normally distributed, and the chip counts from the Brand B cookies should be approximately Normally distributed. Note that parts (b) and (c) ask the students to plot histograms of the sample data from Brand A and Brand B cookies to see if they feel the Normality assumption is being violated. Students have been told that nonparametric tests exist if this Normality assumption is violated, but these have not been discussed yet. It turned out that both histograms appeared to be symmetrical and bell-shaped.

The independent two-sample *t*-test also assumes equal variances (or standard deviations). The lab sheet requires that Minitab be used to calculate the sample standard deviations of the number of chips in each cookie to get a feel for whether this assumption is being violated. In this case, it turned out that the sample standard deviations of the number of chips in Brand A and Brand B cookies were 4.375 and 4.485, respectively. It was not felt that the equal variance assumption was violated.

The null hypothesis was rejected based on the data collected in class. A 95% confidence interval for $\mu_A - \mu_B$ was given by (-7.65, -2.26). Students were able to write a written interpretation for this interval and that it was consistent with rejecting the null hypothesis that the average number of chips in a Brand A cookie was the same as the average number of chips in a Brand B cookie.

Lab Sheet 1

Test: H_0: $\mu_A = \mu_B$ against H_1: $\mu_A \neq \mu_B$ with α = 0.05

Put observations for A in C1 and observations for B in C2.

a) Find the sample means and sample standard deviations of the number of chocolate chips per cookie for both brands of cookies.

$\overline{X}_A$ = ...S_A = ...$\overline{X}_B$ = ...S_B = ...

b) Does the number of chocolate chips in each Brand A cookie appear to be approximately Normally distributed? Graph a histogram. Comment.

c) Does the number of chocolate chips in each Brand B cookie appear to be approximately Normally distributed? Graph a histogram. Comment.

d) Conduct the test. (It will be a 2-sample *t*-test assuming equal variances.)

What is the value of the test-statistic?
What is the *p*-value?
Should H_0 be rejected?
The 95% confidence interval for $\mu_1 - \mu_2$ is........:
Interpret.

Is the result of the test consistent with the result of the confidence interval?_______________________

NOTE: (b) and (c) were added so that students were reminded of one of the assumptions for the test.

• CLASS ACTIVITY TWO •

Class Activity Two involved students collecting data on the taste ratings of both brands of cookies. Each student, if they wanted to, was to taste both brands of cookies and assign them ratings. Everyone wanted to taste and rate the cookies!! The reason for doing this activity was that I had introduced independent two-sample *t*-tests and paired *t*-tests, and I was trying to get students to see the difference between having two independent samples and a paired sample.

The activity involved transforming a qualitative random variable into a quantitative one. Because of this, I had the students in class assume they represented a random sample of professional taste testers. Professional taste testers should have in mind what they are looking for in a cookie and how many points to deduct if something isn't there.

Since one cookie rating in each of the two sets was assigned by the same taste tester, these ratings were dependent, resulting in paired data. It was mentioned to the students that an experiment involving paired data is used when it is thought that pairing ahead of time will significantly reduce sources of variation. In this case, each taste tester would have his/her own ideas about what makes a good cookie, and also how many points to deduct for each missing characteristic. If different taste testers had been used to rate each brand of cookie, more sources of variation would have been introduced.

The following sheet was given to students to complete (Example 2).

> <u>Example 2</u>:
> Assume the students in the class represent a random sample of professional taste testers. (In this case we are testing cookies.) Each student will rate two brands of cookies on a scale of "0"= terrible to "10"= excellent, with "5" = OK. Use whole numbers only.
>
> (Data will be in pairs. Record the rating each student gave the Brand A cookie first, and the Brand B cookie second.)
>
> NOTE: The data actually collected in class were as follows:
>
> *(5,7)* (4,6) (5,8) (5,7) (6,8) (3,9) (5,10) (4,7) (5,7) (5,7) (4,7) (5,8) (6,5) (7,9) (5,3) (1,5) (4,8) *(5,7)* (3,7) (8,7) (4,6) (3,6)

Students were asked not to consult each other before writing down their ratings. Half of the class was given Brand A cookies to taste first, and the other half was given Brand B. Labels had been taken off all cookies ahead of time. Students were asked to take a drink of water between taste tests. The students later took the data to the computer laboratory and completed Lab Sheet 2. Calculations were done using Minitab. In this case, a paired *t*-test was conducted.

> <u>Lab Sheet 2</u>
> Let μ_D = mean difference in ratings between Brand A cookies and Brand B cookies.
> Test: $H_0 : \mu_D = 0$ against $H_1 : \mu_D \neq 0$ with $\alpha = 0.05$.
> Enter observations for Brand A cookies in C1 and corresponding observations for Brand B cookies in C2. You must first compute the differences. Enter these in C3.
>
> a) Find the sample mean and sample standard deviation of the difference in Brand A and Brand B cookie ratings.
> $\overline{X}_D$ = ...S_D = ...
> b) Do the differences appear to be approximately Normally distributed?
> Graph a histogram.
> Comment.
> (Note: In this case, the histogram did appear to be symmetrical and bell-shaped.)
> c) Conduct the test. What is the value of the test-statistic?.....................
>
> What is the p-value?.
> ....
> Conclusion:..
> d) Find a 95% confidence interval for μ_D. The 95% confidence interval for μ_D is:....................
> Interpret. Is this consistent with the test result?

The null hypothesis was rejected based on the sample data. The 95% confidence interval for μ_D was found to be (-3.19, -1.54). Students were again able to interpret this and relate it to rejection of the null hypothesis.

◆DISCUSSION ◆

Students seemed to have a great time collecting both sets of cookie data in class. These were good activities to do because every student related to them, even though they came from a wide variety of backgrounds. Students worked hard in the computer laboratory. They were able to write written interpretations of the results. Namely, instead of just saying H_0 is rejected, they were able to write H_0 is rejected and what that implied.