Double Stuffed?

Allison Dorko
Oregon State University
dorkoa@onid.oregonstate.edu

Overview of Lesson
In this activity students will use statistics to determine whether there is evidence that Double Stuf Oreos have double the filling as Regular Oreos. Students will use two-sample t-tests to determine: (a) if there is evidence that Double Stuf Oreos and Regular Oreos have different amounts of frosting, and (b) if there is evidence that the mean amount of filling in Double Stuf Oreos is twice the mean amount of frosting in Regular Oreos.

Every step of a hypothesis test will be performed including: constructing the hypotheses, checking the necessary assumptions, calculating the test statistic and p-value, and interpreting the results in context.

GAISE Components
This investigation follows the four components of statistical problem solving put forth in the *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report*. The four components are: formulate a question, design and implement a plan to collect data, analyze the data by measures and graphs, and interpret the results in the context of the original question. This is a GAISE Level C activity.

Common Core State Standards for Mathematical Practice
1. Make sense of problems and persevere in solving them.
2. Reason abstractly and quantitatively.
3. Construct viable arguments and critique the reasoning of others.
4. Model with mathematics.
5. Use appropriate tools strategically.
6. Attend to precision.

Common Core State Standards Grade Level Content (High School)
S-ID. 1. Represent data with plots on the real number line (dot plots, histograms, and box plots).
S-ID. 2. Use statistics appropriate to the shape of the data distribution to compare center (median, mean) and spread (interquartile range, standard deviation) of two or more different data sets.
S-ID. 3. Interpret differences in shape, center, and spread in the context of the data sets, accounting for possible effects of extreme data points (outliers).
S-IC. 1. Understand statistics as a process for making inferences about population parameters based on a random sample from that population.
S-IC. 5. Use data from a randomized experiment to compare two treatments; use simulations to decide if differences between parameters are significant.

NCTM Principles and Standards for School Mathematics
Data Analysis and Probability Standards for Grades 9-12

Formulate questions that can be addressed with data and collect, organize, and display relevant data to answer them:

- understand the meaning of measurement data and categorical data, of univariate and bivariate data, and of the term variable;
- understand histograms, parallel box plots, and scatterplots and use them to display data;
- compute basic statistics and understand the distinction between a statistic and parameter.

Select and use appropriate statistical methods to analyze data:

- for univariate measurement data, be able to display the distribution, describe its shape, and select and calculate summary statistics.

Develop and evaluate inferences and predictions that are based on data:

- understand how sample statistics reflect the values of population parameters and use sampling distributions as the basis for informal inference.

Prerequisites

As written, this activity is meant to follow instruction about a two-sample t-test and, as an extension, confidence intervals. It is assumed that students are familiar with calculating measures of central tendency and spread, the normal model, formulating hypothesis statements, and performing t-tests for independent samples.

Learning Targets

After completing the activity, students will be able to identify: (i) when a t-test for two independent samples is appropriate and state the hypotheses, (ii) state and check the assumptions, (iii) carry out the test, and (iv) use the test statistic to justify their conclusion about the questions of interest.

Time Required

The time required for this activity is roughly 90 minutes. The data collection phase is the longest part and can be shortened by acquiring a balance for every 2 students.

Materials Required

For this activity students will need

- Scales (such as a kitchen scale or a triple-beam balance)
- At least one package each of regular and Double Stuf Oreos (for the class)
- Muffin papers (for use as a container to mass frosting)
- Bowls in which to deposit cookies and frosting
- Graphing calculators or statistical computing software
- A copy of the Activity Sheet (attached to the end of this lesson)

Instructional Lesson Plan

The GAISE Statistical Problem-Solving Procedure

I. Formulate Question(s)

Contact Author for permission to use materials from this STEW lesson in a publication

Begin the activity by showing students the packages of Regular and Double Stuf Oreos and ask if they think Double Stuf Oreos have more filling than Regular Oreos. Ask how they could represent this question in the language of a hypothesis test. For the remainder of this activity, let μ_R denote the mean amount of filling (grams) of Regular Oreos and μ_D denote the mean amount of filling (grams) of Double Stuf Oreos. Since it is reasonable to assume that Double Stuf Oreos do indeed have more filling, the hypotheses should be written as: H_0: $\mu_{D-}\mu_R = 0$ vs H_A: $\mu_D - \mu_R > 0$.

Next, ask students if they think Double Stuf Oreos have twice the filling of Regular Oreos. Ask how they could investigate this question statistically. In practice, we will multiply the data from the Regular Oreos by two to test: H_0: $\mu_D - 2\mu_R = 0$ vs H_A: $\mu_D - 2\mu_R \neq 0$.

NOTE: This may be easier for students to think about *after* collecting and analyzing data for the first question.

II. Design and Implement a Plan to Collect Data
Tell students that they have at their disposal packages of each type of Oreos, balances, plastic cutlery, and muffin papers. Have students formulate a procedure for collecting data. The procedure should look something like the following:
   a. Students divide the cookies amongst themselves; each student should find the mass of frosting from both types of cookies. The class should have data from a total of at least 30 of each type of Oreos.
   b. Place the muffin paper on the scale and use the tare feature such that the scale is "tared" to zero.
   c. Scrape the frosting from an Oreo and put the frosting on the muffin paper. Place the muffin paper and frosting on the scale. Read the mass and record it in a class data table.
   d. Repeat steps b and c until the class has collectively found the masses of the frosting from 30 Double Stuf Oreos and 30 Regular Oreos.

To illustrate the data analysis portion of this activity, filling masses from 30 Regular and 30 Double Stuf Oreos are given in Table 1.

Table 1. Example frosting mass data.

Filling Masses of Regular Oreos (g)					
3.00	2.90	3.20	3.20	3.20	3.12
3.15	3.15	3.01	3.19	3.01	3.17
3.06	3.14	3.16	3.31	3.13	3.23
2.94	3.05	3.14	3.39	3.30	3.09
3.10	3.11	3.24	3.18	3.10	3.15
Filling Masses of Double Stuf Oreos (g)					
6.11	6.19	6.27	6.18	6.34	6.21
6.05	6.18	6.34	6.23	6.08	6.04
6.13	6.65	6.20	6.12	6.46	6.37
6.19	6.01	6.21	6.27	6.46	6.15
6.15	6.35	6.24	6.25	6.32	6.03

Contact **Author** for permission to use materials from this STEW lesson in a publication

III. Analyze the Data

Students should check the assumptions. The cookies are independent random samples. Students should create a dot plot, histogram, or boxplot to check for normal distribution and homogeneity of variance. The histograms based on the data above are shown below.
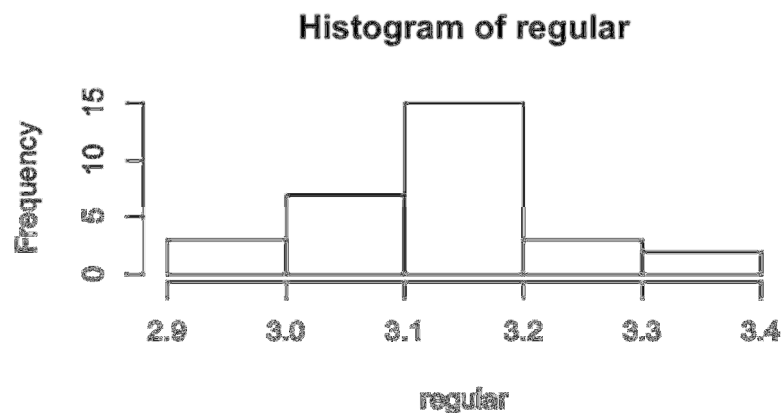
Histogram of regular

Figure 1. Masses of filling in Regular Oreos (g).
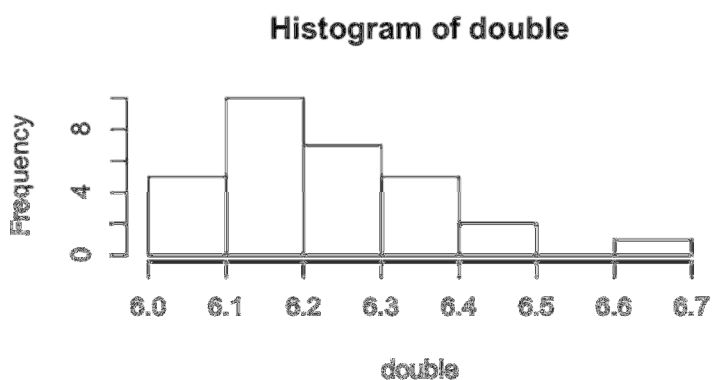
Histogram of double

Figure 2. Masses of filling in Double Stuf Oreos (g).

Students should then specify an alpha level and use their calculator or statistical software to compute the t-statistic. The one-tailed test for H_0: $\mu_{D} - \mu_R = 0$ versus H_A: $\mu_D - \mu_R > 0$ is:

t = 95.2763, df = 58, p-value ≈ 0
mean of Double Stuf Oreos = 6.226g
mean of Regular Oreos = 3.137g

Students should interpret the p-value from the H_0: $\mu_{D} - \mu_R = 0$ vs H_A: $\mu_D - \mu_R > 0$ test before performing the second test. That is, students should conclude whether or not there is evidence that Double Stuf Oreos and Regular Oreos have different amounts of frosting. (Based on the sample data in Table 1, we would conclude that the p-value of approximately zero indicates that there is sufficient evidence that Double Stuf Oreos and Regular Oreos have different amounts of frosting).

The next question of interest is whether or not Double Stuf Oreos have twice the filling of Regular Oreos. Students should brainstorm ways they could test this. We multiply all of the filling masses from Regular Oreos by 2, then run a two-tailed t-test (α=0.05) for a difference in population means. The test statistic for the data above is:
t = -1.0452, df = 58, p-value = 0.3003

Students should then interpret the result from the H_0: $\mu_D - 2\mu_R = 0$ vs. H_A: $\mu_D - 2\mu_R \neq 0$ test. That is, students should conclude whether or not there is evidence that Double Stuf Oreos do not have double the frosting of Regular Oreos. (Based on the sample data in Table 1, we would conclude that the p-value of 0.3003 indicates that there is not sufficient evidence to say that Double Stuf Oreos do not have double the frosting of Regular Oreos).

IV. Interpret the Results
Students should use the t-statistic and p-value from their computations to conclude whether or not: (a) Double Stuf Oreos have more filling than Regular Oreos and (b) whether or not Double Stuf Oreos have twice the filling as Regular Oreos. The conclusions from the sample data above are:
(a) Based on p ≈ 0, students should reject the hypothesis that Double Stuf Oreos have the same amount of filling as Regular Oreos. The one-tailed test that $\mu_D - \mu_R > 0$ indicates that there is sufficient evidence to conclude that Double Stuf Oreos have more filling than Regular Oreos.
(b) Based on the p-value of 0.3003, we fail to reject the null hypothesis that $\mu_D = 2\mu_R$. In other words, we fail to reject that Double Stuf Oreos have twice the filling of Regular Oreos. In practice, we cannot conclude that on average, Double Stuf Oreos do not have twice the filling of Regular Oreos.

Assessment

The following questions assess what students learned in the lesson by asking students to use the same tests in a different context. We thus assess that students can identify the correct test (i), state the hypotheses (i), state and check the necessary assumptions (ii), perform the test (iii), and interpret it in context (iv). As in the lesson, students need to think about what to do with their data such that they can test whether one mean is double another. The Roman numerals in the assessment problem correspond to the Roman numerals in the Learning Targets section.

The data below are the salaries for 14 randomly selected teachers and 16 randomly selected doctors. For the two questions below,

i.   State the test you will use, whether or not it is one-tailed or two-tailed, and its hypotheses. Test at $\alpha = 0.05$.

ii.   Check the assumptions.

iii.   Calculate the test statistic (you may use statistical computing software).

iv.   Interpret the test statistic in context and use it to answer the question regarding whether or not the salaries differ.

Teachers' Salaries (USD)		Doctors' Salaries (USD)	
53118	54963	112184	105406
55740	58627	106919	110197
62899	60448	110176	109944
58500	61842	111740	113950
56726	56873	115535	104116
55554	64103	110108	116054
59247	61138	114815	105119
		108114	108549

1. Do doctors make more money than teachers?

2. Do doctors make twice as much as teachers?

Contact Author for permission to use materials from this STEW lesson in a publication

Answer Key

Question 1

i. Two-sample t-test for independent samples (one-tailed).

H_0: $\mu_{D-\mu_T} = 0$ ($\mu_D = \mu_T$)

H_A: $\mu_D - \mu_T > 0$

ii. The data come from random samples. Teachers' and doctors' salaries are independent (we can assume no teachers also work as doctors, and vice versa). There are no obvious outliers, and the histograms (see below) are mound-shaped and roughly symmetric, so there is no significant reason to question the condition of approximate normality. The salaries also seem to meet the homogeneity of variance assumption.

Histogram of teach
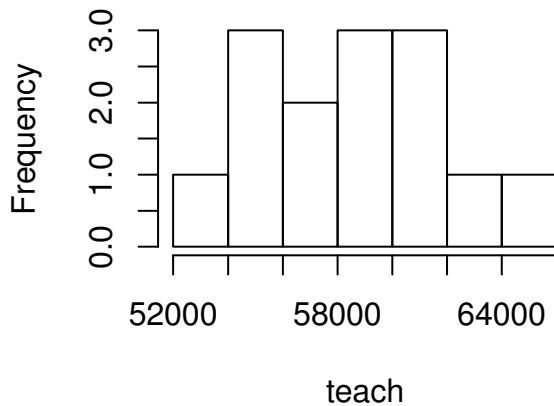
Histogram of doc
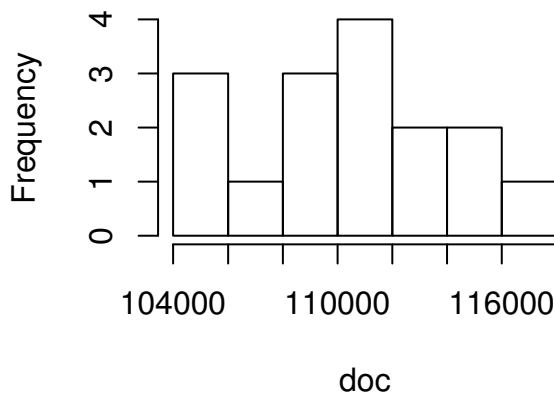
iii. t = 40.1958, df=28, p-value ≈ 0

mean salary for doctors $110,182.88

mean salary for teachers $58,555.57

---

iv. Since p-value ≈ 0 is less than the alpha level of 0.05, we reject the null hypothesis that doctors' and teachers' salaries are equal and conclude that doctors' salaries are higher than teachers'.

Question 2

   i. Two-sample t-test for independent samples (two-tailed)

     H_0: $\mu_D - 2\mu_T = 0$ ($\mu_D = 2\mu_T$)

     H_A: $\mu_D - 2\mu_T \neq 0$

  ii. The assumptions are still met (multiplying the teachers' salaries by 2 is multiplication by a scalar, which does not change the shape of the distributions).

 iii. $t = -3.6454$, df = 28, p-value = 0.001078

 iv. Since $p = 0.001078$ is less than the alpha level of 0.05, we reject the null hypothesis that doctors' salaries are double teachers' salaries. We conclude that doctors' salaries are either *more than* double teachers' salaries, or *less than* double teachers' salaries, though we cannot be sure from this test which of those two cases is true.

Possible Extensions

One extension of this activity is to compute a confidence interval for the mass of frosting in the Regular Oreos and a confidence interval for the mass of frosting in the Double Stuf Oreos.

References

1. Activity background adapted from http://www.cnn.com/2013/08/21/us/oreo-high-school-experiment/

2. American Statistical Association (2007). Guidelines for assessment and instruction in statistics education. Available online: http://www.amstat.org/education/gaise/

Double Stuffed? Activity Sheet

1a. Do you think that Double Stuf Oreos have more filling than Regular Oreos? How could you express this in the language of hypothesis testing?

1b. What sort of statistical test will you use to answer your question? What will you use for an alpha level?

1c. Talk with your classmates and devise a plan to collect data with which to test your hypothesis.

1d. Collect your data and record it in the table below.

Filling Masses of Regular Oreos (g)					
Filling Masses of Double Stuf Oreos (g)					

2a. What are the assumptions underlying your test? Check these assumptions.

2b. Perform your test (you may use statistical software).

2c. Interpret your result. (What does your result say about the amount of filling in Regular and Double Stuf Oreos?)

3a. Do you think that Double Stuf Oreos have twice the filling of Regular Oreos? How could you test this using your data?

3b. Carry out your test.

3c. Interpret your result. (What does your result say about the amount of filling in Double Stuf Oreos: is there evidence to conclude that the amount of filling in Double Stuf Oreos is not double that in Regular Oreos?)

4. More generally, what did you learn about determining whether the amounts of filling in Regular and Double Stuf Oreos differ?