10,000 Steps?

Phyllis Curtis
Grand Valley State University
curtissp@gvsu.edu

Mary Richardson
Grand Valley State University
richamar@gvsu.edu

Published: May 2014

Overview of Lesson
This activity introduces students to the use of pedometers. Pedometers measure the up and down motion of a person's hip in a vertical plane. It is actually said that they were invented by Leonardo da Vinci. They have been used for centuries since people have been interested in measuring the distance they travel by walking and/or jogging. (Pangrazi et al. 2003)

Students hear about the benefits of using a pedometer to track the number of steps they take each day during normal activity, and realize that fitness guidelines suggest walking 10,000 steps per day (www.thewalkingsite.com), which is close to 5 miles.

Students will walk 264 feet, one-twentieth of a mile, while wearing a pedometer. They will investigate if there is a significant difference between the number of steps recorded by a Dollar Store pedometer and a pedometer App.

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.
4. Model with mathematics.
5. Use appropriate tools strategically.

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-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:**
- compute basic statistics and understand the distinction between a statistic and a 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
Students will have knowledge of calculating numerical summaries for one variable (mean, median, five-number summary, checking for outliers).
Students will have knowledge of how to construct boxplots.
Students will have been exposed to the steps in a hypothesis test and will have previously been exposed to confidence intervals.

Learning Targets
Students will be able to use calculated numerical summaries to describe a data set.
Students will understand the distinction between paired samples and independent samples.

Time Required
1 class period (to complete the lesson).

Materials Required
Activity Sheet (see the end of the lesson); pencil and paper; graphing calculator; statistical software package (optional); one Dollar Store pedometer per student or enough of them for students to use on a rotational basis; access to a smart phone pedometer App for each student.

Note: Not all students may have a smart phone, in which case we would suggest having the students count the actual number of steps taken to compare against the Dollar Store pedometer count. A paired t-test can then be done on the data using the actual step count in place of the pedometer App step count. Refer to the extension (below) for more guidance.

Instructional Lesson Plan

The GAISE Statistical Problem-Solving Procedure

I. Formulate Question(s)
According to http://www.livescience.com/43956-walking-10000-steps-healthy.html if you buy a smart pedometer or fitness tracker like a Fitbit, chances are the device will encourage you to take 10,000 steps a day.

The origin of the 10,000 steps recommendation isn't exactly scientific. Pedometers sold in Japan in the 1960s were marketed under the name "manpo-kei," which translates to "10,000 steps meter," said Catrine Tudor-Locke, director of the Walking Behavior Laboratory at Pennington Biomedical Research Center in Baton Rouge, La. The idea resonated with people, and gained popularity with Japanese walking groups, Tudor-Locke said.

Studies conducted since then suggest that people who increased their walking to 10,000 steps daily experience health benefits. One study found that women who increased their step count to nearly 10,000 steps a day reduced their blood pressure after 24 weeks. Another study of overweight women found that walking 10,000 steps a day improved their glucose levels.

However, walking 10,000 steps a day is not an official recommendation from the Centers for Disease Control and Prevention. Instead, the agency recommends adults engage in 150 minutes of moderate activity a week, such as brisk walking. To meet the CDC's recommendation, you would need to walk about 7,000 to 8,000 steps a day, Tudor Locke said.

After discussing with students the idea of having a goal of 10,000 steps per day for fitness, introduce the question of interest: If a student walks one-twentieth of a mile, will there be a significant difference between the number of steps recorded by two different types of pedometers?

II. Design and Implement a Plan to Collect the Data
The teacher should distribute an Activity Sheet and a Dollar Store pedometer to each student in the class. Students should be instructed to clip the pedometer on the right-side of their waist just over the knee. Additionally, students should be sure to reset the pedometer so it reads zero prior to proceeding with the experiment.

Those students who have smart phones should be asked to download a (free) pedometer App and to determine how the App works. One-twentieth of a mile, or 264 feet, should be pre-measured in a hallway near the classroom. Tell students to walk along the pre-determined 264 foot course. Tell them to walk as naturally as possible. At the finish line, they should check the Dollar Store pedometer to see how many steps were recorded. Those students using a smart phone pedometer App should also note the number of steps recorded by the App. Then, they should proceed back to the classroom and write the Dollar Store pedometer reading along with the pedometer App reading (if appropriate) on the board in the specified location.

Note: Several free pedometer Apps are available for smart phone users. Pedometer Free GPS+ by Arawella Corporation was used in this classroom. This pedometer App was recommended at the following website out of four Apps tested: http://www.discoverwalking.com/blog/best-iphone-free-pedometer-App.php.

III. Analyze the Data
This experiment was piloted in a small undergraduate statistics major course and a relatively small sample size was obtained due to not every student having a smart phone. The example class data is provided in the following table.

Table 1. Example class data.

Student	Dollar Store Pedometer Steps	Pedometer App Steps
1	200	222
2	201	208
3	217	174
4	173	124
5	201	173
6	225	210
7	209	123
8	327	315
9	174	123
10	199	208

The teacher can begin the data analysis by asking students to explain why the two samples, the number of steps recorded by the Dollar Store pedometer and the pedometer App, cannot be treated as independent samples. Students should note that since the same person used both types of pedometers, that this data is matched (or paired). Then, ask students to explain how the data collection process could be changed in order to obtain independent samples to compare the average steps recorded for the two types of pedometers. In order to obtain independent samples, each person should wear only one type of pedometer with the type being decided by flipping a coin.

Ask: "Which data collection method would you recommend for this activity – independent or paired?   Why?" Students should answer paired since the variability from person to person will be reduced by using a paired data collection design. A paired design is preferable to an independent samples design here because the pairing helps to control for differences in step counts across students. If a student obtains different step counts for the two types of pedometers we feel more confident attributing the difference to the type of pedometer than we would if we compared two different students.

Have students calculate the difference between the number of step counts for the Dollar Store pedometer and the pedometer App (Dollar Store minus App).

Table 2. Differences for the example class data.

Student	Dollar Pedometer Steps	Pedometer App Steps	Difference in Steps Recorded
1	200	222	-22
2	201	208	-7
3	217	174	43
4	173	124	49
5	201	173	28
6	225	210	15
7	209	123	86
8	327	315	12
9	174	123	51
10	199	208	-9

Discuss with students that in order to perform a paired t-test using this data, we need to be able to assume that the sampling distribution of the paired differences is approximately normal. To check this assumption, we should focus on the differences since the assumptions for paired tests are based upon the differences.

Students calculate appropriate summary measures and construct a stemplot for the differences. The summary calculations and the stemplot are shown below.

Minimum	-22
Maximum	86
Median	21.5
Mean	24.6
Standard Deviation	33.3

```
Stem Leaf
  8 6
  6
  4 391
  2 8
  0 25
 -0 97
 -2 2
```
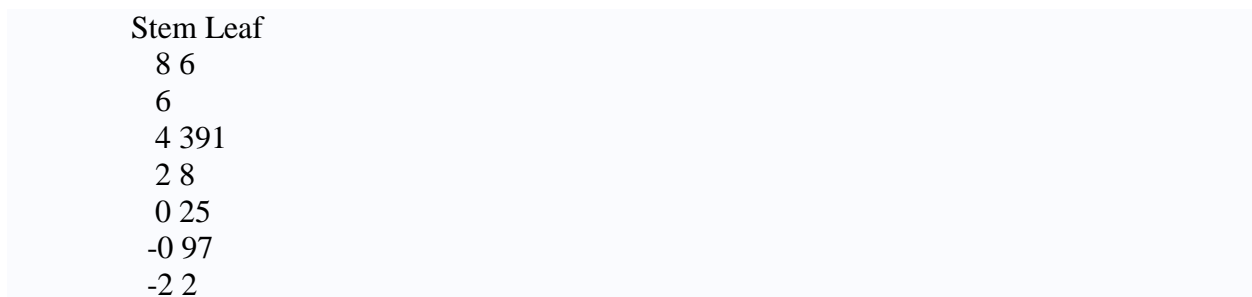
Figure 1. Summary calculations and stem plot of the differences.

Based on the stemplot students are asked if they think the sampling distribution of the paired differences is approximately normal. Students should note that the stemplot is somewhat normal shaped and there do not seem to be any outliers.

Next, the mean and median differences are compared. Recall that if the distribution of the differences fits a normal distribution, the mean and the median will be roughly the same. For this distribution, the mean is 24.6 and the median is 21.5. These two values are close, especially since the standard deviation is 33.3.

Lastly, based upon their checks, students are asked if they think that we can assume the sampling distribution of the paired differences is approximately normal. We can assume normality since none of the checks showed evidence against it.

IV. Interpret the Results

Students obtain a boxplot of the differences in step counts and are asked "If it is actually the case that the two types of pedometers count steps differently, what features should be seen in the difference data distribution?" Students should note we should expect to see a large percentage of either positive or negative differences. A boxplot for the example class difference distribution is shown below.

```
          Boxplot
 8           |
 6           |
 4        +-----+
 2        *--+--*
 0        |   |
-0        +-----+
-2           |
```

Figure 2. Boxplot of class differences.
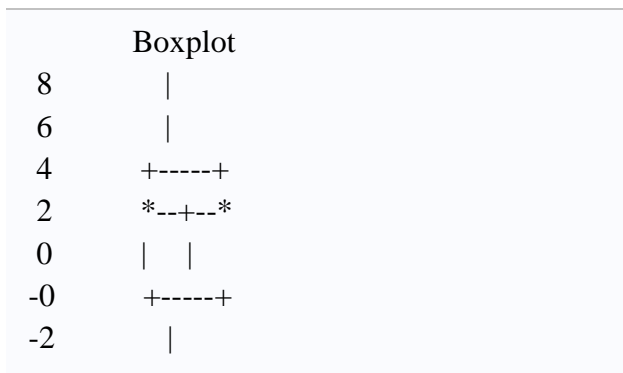
Students determine whether the boxplot provides evidence in either direction regarding a difference in step counts for the two pedometer types. They state the aspects of the boxplot that can be used to justify their answer. From the difference boxplot, we can see that over half of the differences are positive, this indicates that there may be a difference in the average steps.

Now a hypothesis test is performed to decide if there is a significant difference in the average number of steps recorded for the two types of pedometers. The hypotheses are: H$_0$: $\mu_d = 0$ versus H$_A$: $\mu_d \neq 0$ where μ_d is the mean difference in number of steps counted for the Dollar Store pedometer minus the App pedometer. The *t* test statistic is equal to 2.39 and it is based upon 9 degrees of freedom. Note: the formula used for the test statistic is: $t = \dfrac{\bar{d} - 0}{\frac{s_d}{\sqrt{n}}}$ with $\bar{d}$

being the average sample difference and s_d being the standard deviation of the sample differences. The p-value is .0442. Thus, based upon a .05 level of significance, the null hypothesis will be rejected and we will conclude that there is significant evidence at the 5% level to show there is a difference in the mean number of steps recorded by the Dollar Store pedometer and the App pedometer. The Dollar Store pedometer recorded more steps.

A 95% confidence interval for the mean difference in the number of steps recorded for the two different types of pedometers is given by: (.80, 48.40). Thus, we are 95% confident that the mean number of steps recorded by the Dollar Store pedometer is between .80 and 48.4 steps larger than the mean number of steps recorded by the pedometer App.

A discussion on how the confidence interval gives the same conclusion as the hypothesis test concludes the activity. Students should note that both of the confidence interval endpoints are positive indicating that there is a difference in the means.

Assessment

1. In each part, identify whether it would be appropriate to use a procedure for paired data or a procedure for two independent samples to estimate the difference of interest.

(a) A researcher wants to estimate the difference between the mean scores on a memorization test for people over 60 years old compared to people under 40 years old.

(b) Women give their actual and desired weights in a survey. An estimate of the mean difference is desired.

(c) Two treatments for reducing cholesterol are compared. Forty people with high cholesterol use the first treatment, and 40 other people with high cholesterol use the second treatment. An estimate of the difference between the mean decrease in cholesterol for treatment 1 and the mean decrease in cholesterol for treatment 2 is desired.

2. A random sample of five college women was asked for their own heights and their mothers' heights. The researchers wanted to know whether college women are taller on average than their mothers. The results (in inches) were as follows:

Pair	1	2	3	4	5
Daughter	66	64	64	69	66
Mother	66	62	65	66	63

(a) Define the parameter of interest in this situation.

(b) Find a 95% confidence interval for the parameter you defined in part (a).

(c) Using the interval in part (b), write a sentence or two about the relationship between women students' heights and their mothers' heights for the population. Your explanation should be written to be understood by someone with no training in statistics.

(d) Explain two different things that the researchers could have done to obtain a narrower interval than the one you found in part (b).

3. Give the value of the test statistic t in each of the following situations, then find the p-value for a two-tailed test.

(a) $H_0 : \mu_d = 0, \ \bar{d} = 4, \ s_d = 15, \ n = 50.$

(b) $H_0 : \mu_d = 0, \ \bar{d} = -4, \ s_d = 15, \ n = 50.$

(c) $H_0 : \mu_d = 0, \ \bar{d} = 0, \ s_d = 15, \ n = 50.$

4. A study was reported in which students were asked to place as many dried beans into a cup as possible in 15 seconds with their dominant hand, and again with their non-dominant hand (in randomized order). The differences in number of beans (dominant hand minus non-dominant hand) for 15 students were as follows: 4, 4, 5, 1, -2, 0, 2, 4, -3, 0, 0, 0, -2, 2, 1.
(a) The research question was whether students have better manual dexterity with their dominant hand than with their non-dominant hand. Write the null and alternative hypotheses.

(b) Check the necessary conditions for doing a paired t-test.

(c) Carry out the test using a .05 level of significance.

(d) Carry out the test using a .10 level of significance.

(e) Write a conclusion about this situation that would be understood by other students of statistics.

Answers

1. (a) Procedure for two independent samples.
(b) Procedure for paired data.
(c) Procedure for two independent samples.

2. (a) The parameter of interest is μ_d = the mean difference in heights (in inches) between college women and their mothers.

(b) Compute interval as Sample estimate ± Multiplier × Standard error, which here is $\bar{d} \pm t^* \dfrac{s_d}{\sqrt{n}}$.

Sample estimate is $\bar{d} = 1.4$. Standard error is s.e. $(\bar{d}) = \dfrac{s_d}{\sqrt{n}} = \dfrac{1.817}{\sqrt{5}} = 0.813$.

Multiplier is $t^* = 2.78$; use df $= n - 1 = 5 - 1 = 4$.
The interval is $1.4 \pm (2.78 \times 0.813)$, or 1.4 ± 2.26, or -0.86 to 3.66.
(c) With 95% confidence we can say that college women are somewhere between 0.86 inch shorter and 3.66 inches taller than their mothers, on average. Or, in other words, we cannot conclude that there is a significant different in the average heights of mothers and daughters.
(d) They could have used a larger sample size, which would have reduced both the standard error and the multiplier (with the same confidence level). They could also use a lower confidence level; for instance they could find a 90% confidence interval instead of a 95% interval.

3. (a) $t = \dfrac{(4-0)}{15/\sqrt{50}} = 1.89$, df $= 50 - 1 = 49$, p-value is between $2 \times .04 = .08$ and $2 \times .026 = .052$.

(b) $t = \dfrac{(-4-0)}{15/\sqrt{50}} = -1.89$, df $= 49$ and p-value is the same as in part (a).

(c) $t = \dfrac{(0-0)}{15/\sqrt{50}} = 0$; df $= 49$ and p-value $= 2 \times .5 = 1.0$.

4. (a) $H_0: \mu_d = 0$ (could also be written as $H_0: \mu_d \leq 0$)

$H_a: \mu_d > 0$ (Manual dexterity is better for dominant hand, on average)

μ_d = mean "dominant hand–nondominant hand" difference in beans placed for population represented by the sample
(b) A dotplot and/or a boxplot can be used to verify that the sample of differences is more or less symmetric and there are no outliers. It is assumed the sample represents a random sample from a larger population.
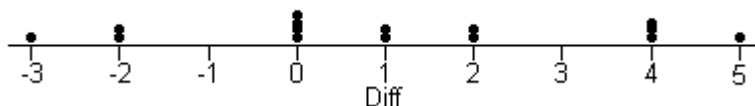
(c) Test statistic is $t = \dfrac{\text{Sample statistic-Null value}}{\text{Null standard error}} = \dfrac{1.067 - 0}{0.628} = 1.70$

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

Sample statistic is $\bar{d} = 1.067$; null standard error $= \dfrac{s_d}{\sqrt{n}} = \dfrac{2.434}{\sqrt{15}} = 0.628$

Minitab output follows:

```
Test of mu = 0 vs mu > 0

Variable       N    Mean    StDev   SE Mean
Diff          15    1.067   2.434   0.628


Variable    95.0% Lower Bound     T     P
Diff              -0.040       1.70  0.056
```

p-value = .056. Do not reject the null hypothesis for α = .05. With this significance level, we are not able to conclude that, on average, the number of beans placed is greater for the dominant hand.

Note: The p-value is the area (probability) to the right of t = 1.70 in a t-distribution with df = n−1= 15−1 =14.

(d) Reject the null hypothesis for α =.10 (because p-value = .056 < .10). With this level of significance we can say that, on average, individuals in the population can place more beans with the dominant hand.

(e) The p-value (.056) is not quite below the α =.05 standard, but it is less than or equal to any $\alpha \geq .056$. So, statistical significance can be claimed for any significance level of .056 or higher. One difficulty here is the small sample size. A larger sample size would provide more conclusive evidence about the mean difference between beans placed with the two hands. Also, the task is relatively easy, so the mean difference between the hands may not be large and may not be the best way to compare manual dexterity for the two hands.

Possible Extension

In this activity students tested for a difference in the average step counts recorded by two types of pedometers. To extend the analysis, ask students: "Suppose you wanted to test if the average number of steps recorded by the Dollar Store pedometer differs from the actual number of steps taken. How would you collect data to answer this question? What hypothesis test would be the best one to perform (assume any needed conditions are met) in order to answer this question?" One possibility: During the data collection phase, decide on a set distance for the students to walk (for example: one twentieth of a mile). While walking with the Dollar Store pedometer, each student could be instructed to also count the actual number of steps taken. The paired t test could be performed using the Dollar Store pedometer and the actual steps. A second possibility: During the data collection phase, each student could be instructed to take 200 steps and then record the number of steps counted by the Dollar Store pedometer. The one-sample t test could be performed using a null value of 200 steps.

References

1. *Guidelines for Assessment and Instruction in Statistics Education* (GAISE) Report, ASA, Franklin et al., ASA, 2007 http://www.amstat.org/education/gaise/ .

2. Assessment questions taken from: *Mind on Statistics*, 3rd Edition by Utts/Heckard, 2006. Cengage Learning.

3. Pangrazi, R, Beighle, A., and Sidman, C. (2003) *Pedometer Power*, Human Kinetics, Champaign, Il.

10,000 Steps? Activity Sheet

Background: (Information provided is from www.thewalkingsite.com.)
Maybe you have heard the recent guidelines about walking 10,000 steps per day. How far is 10,000 steps anyway? The average person's stride length is approximately 2.5 feet long. That means it takes just over 2,000 steps to walk one mile, and 10,000 steps is close to 5 miles. Do you need a little motivation to get out the door? Try wearing a pedometer and gradually increasing your daily steps. A pedometer will keep track of the number of steps you take, but just how accurate is it? If the pedometer says 100 steps, did you really take exactly 100 steps?

Problem:
If a student walks one-twentieth of a mile, will there be a significant difference between the number of steps recorded for two types of pedometers?

Instructions:
If you have a smart phone, download a (free) pedometer App and experiment with using the App. Dollar Store pedometers will be distributed to every student in the class. When you get your Dollar Store pedometer, you should clip it on the right-side of your waist just over the knee. Make sure you reset the pedometer so it reads zero. One-twentieth of a mile, or 264 feet, has been measured in the hallway outside the door. As you walk this course, walk as naturally as possible and stop counting when you cross the tape finish line. Check both of your pedometers to see how many steps that were recorded. Proceed back to the classroom when you are finished and write the number of steps you counted and the pedometer readings on the board in the specified location.

Data collection: Each of you used a Dollar Store pedometer and a pedometer App. You walked one-twentieth of a mile and recorded the number of steps counted by the pedometers. Copy the class data into the table below.

Student	Dollar Store Pedometer Steps	Pedometer App Steps
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		

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

12		
13		
14		
15		
16		
17		
18		
19		
20		

Problem: We want to determine if there is a significant difference in the average steps recorded for two different types of pedometers – in this case we are using a Dollar Store pedometer and a pedometer App.

1. Explain why these two samples, the number of steps recorded by the Dollar Store pedometer and the pedometer App, cannot be treated as independent samples.

2. Explain how the data collection could be changed in order to obtain independent samples to compare the average steps recorded for the two types of pedometers.

3. Which data collection method would you recommend for this activity – independent or paired? Why?

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

4. Calculate the differences in the number of step counts between the Dollar Store pedometer and the pedometer App (Dollar Store minus App):

5. Checking Normality: In order to perform a paired t-test using this data, we need to be able to assume that the sampling distribution of the paired differences is approximately normal.

    (a) To check this assumption, should we obtain a graph of the differences or construct side-by-side graphs of the two step counts? Give a brief explanation of your answer.

    (b) Calculate appropriate summary measures and construct a stemplot for the differences obtained in Question 4.

Minimum	
Maximum	
Median	
Mean	
Standard Deviation	

(c) Based on the stemplot do you think the sampling distribution of the paired differences is approximately normal? Explain, taking into account the shape, sample size and outliers.

(d) Compare the mean and the median difference. Recall that if the distribution of the differences fits a normal distribution, the mean and the median will be roughly the same. Are the mean and the median roughly the same for this distribution? Comment.

(e) Referring to your answers in parts (a) through (d), do you think that we can assume that the sampling distribution of the paired differences is approximately normal?

6. Obtain a boxplot of the differences in step counts. Based on the boxplot, do you think there is a difference in the average steps recorded on the two types of pedometers? Explain.

7. Perform a hypothesis test to decide if there is a significant difference in the average number of steps recorded for the two types of pedometers.

Hypotheses: H_0: H_A:

Test Statistic and degrees of freedom:

P-Value:

Conclusion:

8. Construct a 95% confidence interval for the mean difference in the number of steps recorded for the two different types of pedometers. Interpret the confidence interval.

9. Explain how the confidence interval gives the same conclusion as the hypothesis test.