

An A-MAZE-ING Comparison

Matt Malloure
Grand Valley State
University

matt.malloure@gmail.com

Mary Richardson
Grand Valley State
University

richamar@gvsu.edu

Neal Rogness
Grand Valley State
University

rognessn@gvsu.edu



Published: February 2012

Overview of Lesson Plan

In this activity students will have the opportunity to explore real data collected on the completion of a maze. Through the two-sample independent t -test, students will test to see if the mean time to complete the maze is significantly different for males compared to females. 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 activity follows all 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.

Common Core State Standard 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

Before students begin the activity they will be practiced in calculating descriptive statistics and constructing graphical summaries of univariate data. They will have also learned about sampling distributions for means, constructed confidence intervals for means, and performed hypothesis tests on one mean. The activity will commence after the students are introduced to two-sample t -tests.

Learning Targets

After completing the activity, students will be able to calculate descriptive statistics (mean, standard deviation, five number summary, interquartile range) and use them to compare two data sets. Students will also be able to construct comparative boxplots to compare the distribution of data between two groups. Students will be able to carry out a two-sample t -test and use the results to offer a conclusion to the question of interest.

Time Required

The time required for this activity is roughly 1 class period.

Materials Required

For this activity students will only need to bring a pencil, graphing calculator, and scratch paper. The instructor will provide the activity sheet.

Instructional Lesson Plan

The GAISE Statistical Problem-Solving Procedure

I. Formulate Question(s)

Begin the activity by passing out the Activity Worksheet and explaining that the goal of the activity is to determine if there is a significant difference in the mean number of seconds it takes for males and females to complete a maze. The question of interest for this activity is:

Does the mean time in seconds to complete a maze significantly differ between males and females?

The rest of this investigation will be based on this specific question of interest.

II. Design and Implement a Plan to Collect the Data

The maze is included on the Activity Worksheet and was generated from the web site: <http://hereandabove.com/maze/mazeorig.form.html>. It is important that students not look at the maze before making their attempt to complete it in order to keep their results as unbiased as possible. Also make sure that all students have a clean view of a clock or timer. At an agreed upon time all students will begin the maze and upon completion they will record how long it took them to make it through the maze to the nearest second. Students record their maze completion time value on a post-it note to be collected by the teacher. Class maze completion times are written in an appropriate column on the board (male or female). After all student results have been recorded on the board, students need to write the class data in the “Data:” portion of the Activity Worksheet.

To illustrate the data analysis portion of this activity, maze completion times from two class periods are presented in Table 1.

Table 1. Example class data.

Male Completion Time (in Seconds)		Female Completion Time (in Seconds)		
285	285	120	335	185
251	76	88	172	131
220	54	147	217	277
94	285	270	337	94
75	77	160	178	140
180	257	223	264	177
204	112	80	230	121
188		79	108	364
119		256	205	94
88		97	180	125
103		85	122	
176		337	136	

III. Analyze the Data

To begin the data analysis have students write the hypotheses for the two-sample t -test being carried out in this situation. There are two parameters of interest and students can denote them as either μ_1 and μ_2 , where 1 represents the male group and 2 the female group, or μ_M and μ_F . For the remainder of the activity, all subscripts pertaining to males will be M and females will be F. So the hypotheses should be written as:

$$H_0: \mu_M - \mu_F = 0 (\mu_M = \mu_F) \text{ and } H_a: \mu_M - \mu_F \neq 0 (\mu_M \neq \mu_F).$$

Next have students check assumptions to determine whether the parametric two-sample t -test is appropriate. On the Activity Worksheet there is a blank table and plot where students must calculate the descriptive statistics listed and then create a side-by-side boxplot to visually check the assumptions (in addition to being independent random samples with equal variances the distributions of the maze completion times for males and females should both be approximately normal). Using the data for this activity, the students should produce a completed table and boxplots similar to Table 2 and Figure 1.

Table 2. Descriptive statistics for time to maze completion for males and females.

	Males	Females
Sample Size	19	34
Mean	164.68	180.41
Standard Deviation	81.75	83.38
Minimum	54.00	79.00
1st Quartile	88.00	120.00
Median	176.00	166.00
3rd Quartile	251.00	230.00
Maximum	285.00	364.00

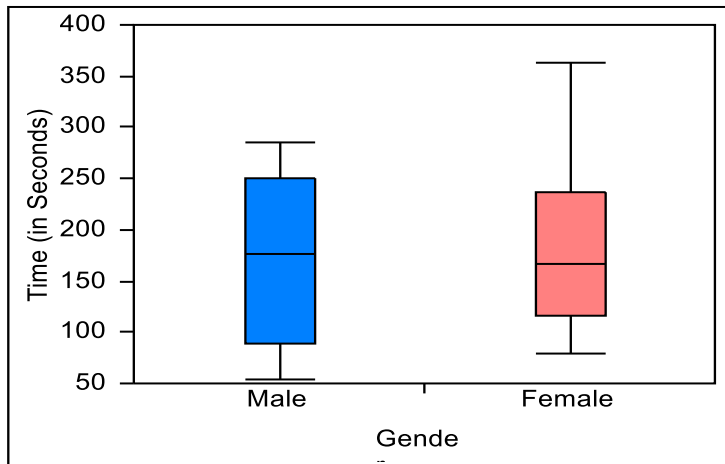


Figure 1. Comparative boxplots for the time to maze completion for males and females.

Students should be able to extend the assumptions for a one-sample t -test to the two-sample t -test. Students should understand that their class isn't exactly a random sample; however, for the purposes of the activity, as long as they make the connection that the class is a sample that may represent some larger population, then this assumption can be considered met. Next, the sample standard deviations for the males and females are quite similar with $s_M = 81.75$ and $s_F = 83.38$. These standard deviations offer little evidence for the two population variances being significantly different. Therefore, students should conclude that $\sigma_M^2 = \sigma_F^2 = \sigma^2$ and the pooled two-sample t -test is appropriate. The final assumption requires students to examine the boxplots in Figure 1. They are looking to check the normality assumption within each group. If the times to maze completion were indeed normally distributed within each group, then the distances between portions of the boxplot should all be roughly equal. There may be some slight skewness in the boxplots presented in Figure 1, but for the most part, the normality assumption appears to be satisfied. The data from Table 1 appear to meet the assumptions required for the two-sample t -test.

The next step in the hypothesis test is to calculate the t test statistic. Since the students determined that a pooled t -test is appropriate, the pooled variance estimate must be calculated prior to the computation of the test statistic. The equation for the estimated pooled variance is:

$$s_p^2 = \frac{s_M^2(n_M - 1) + s_F^2(n_F - 1)}{n_M + n_F - 2}.$$

Note that the estimated pooled standard deviation, s_p , is calculated by taking the square root of the pooled variance estimate ($s_p = \sqrt{s_p^2}$). Using the values from Table 2, students should find the pooled standard deviation estimate to be:

$$s_p = \sqrt{\frac{81.75^2(19 - 1) + 83.38^2(34 - 1)}{19 + 34 - 2}} = 82.81.$$

With the estimated pooled standard deviation in hand, students can now determine the t test statistic, whose formula is:

$$t = \frac{\bar{x}_M - \bar{x}_F}{\sqrt{s_p^2 \left(\frac{1}{n_M} + \frac{1}{n_F} \right)}}$$

Once again, the statistics calculated by the students in Table 2 will provide everything needed to find the t value:

$$t = \frac{164.68 - 180.41}{\sqrt{82.81^2 \left(\frac{1}{19} + \frac{1}{34} \right)}} = -.66.$$

The p-value for this hypothesis test is based on $df = n_M + n_F - 2 = 19 + 34 - 2 = 51$. Students can either use a statistical software package, a statistical table, or their graphing calculator to determine the p-value:

$$P\text{-value} = 2P(t_{51} \leq -.66) = 2(.256) = .51.$$

Students should conclude that at the 5% significance level, the null hypothesis should not be rejected since $.51 > .05$. Their interpretation should say that at the 5% significance level, there is insufficient statistical evidence to indicate that the mean time to maze completion for males and females does significantly differ. Based on the calculations in Table 2 and the boxplots in Figure 1 students may have already suspected that the two means were not going to differ significantly.

Now that the hypothesis test is completed, students should see if a 95% confidence interval for the difference in two independent means produces a similar result. The confidence interval formula that students should implement is:

$$(\bar{x}_M - \bar{x}_F) \pm t^* \sqrt{s_p^2 \left(\frac{1}{n_M} + \frac{1}{n_F} \right)}.$$

The critical value of t^* is such that the area under the t -distribution with $df = n_M + n_F - 2$ between $-t^*$ and t^* is equal to the confidence level. Students may have to use a t critical value table from a textbook to determine the appropriate critical value. For the example data in Table 1, the resulting confidence interval would be as follows:

$$(164.68 - 180.41) \pm 2.01 \sqrt{82.81^2 \left(\frac{1}{19} + \frac{1}{34} \right)} = (-63.41, 31.95).$$

Students should conclude that they can be 95% confident that the true difference in mean time to maze completion for males and females falls between -63.41 and 31.95 seconds. Since 0 is contained in this interval, students should conclude that there is no significant difference between males and females. The confidence interval and hypothesis test both produce the same results.

One important note to consider: if the collected data in another classroom setting shows that the population variances for the two groups may not be equal then the formula for the test statistic in the hypothesis test becomes:

$$t = \frac{\bar{x}_M - \bar{x}_F}{\sqrt{\left(\frac{s_M^2}{n_M} + \frac{s_F^2}{n_F} \right)}}$$

and the confidence interval formula will be:

$$(\bar{x}_M - \bar{x}_F) \pm t^* \sqrt{\left(\frac{s_M^2}{n_M} + \frac{s_F^2}{n_F} \right)}$$

with an approximation for df equal to the smaller of the two sample sizes minus 1.

IV. Interpret the Results

During the activity students were asked to interpret their results from the hypothesis test and confidence interval. For the example presented here students should conclude that there is not enough statistical evidence to indicate that the mean times to maze completion differ between males and females. This result could also be seen from the descriptive statistics in Table 2 and from the boxplots in Figure 1. Some students may make the connection between the confidence interval formula and the calculation of the test statistic, but it is more important for them to understand that the two techniques will result in the same conclusions.

Assessment

1. Each of 63 students in a statistics class used her or his non-dominant hand to print as many letters of the alphabet, in order, as they could in 15 seconds. The following table includes the descriptive statistics for the data collected:

Sex	n	Mean	StDev
Female	29	12.55	4.01
Male	34	13.65	4.46

Complete the following problems to calculate a 95% confidence interval for the difference in the mean number of printed letters between males and females. Assume the population variances are equal (pooled interval) and $t^* = 2.00$.

(a) What is the parameter of interest in this example? What is the sample statistic for this parameter? Calculate its value from the table above.

(b) Calculate the pooled standard deviation estimate.

(c) Determine the 95% confidence interval.

(d) Interpret the interval in the context of the problem.

2. For a sample of 36 men, the mean head circumference is 57.5 cm with a standard deviation equal to 2.4 cm. For a sample of 36 women, the mean head circumference is 55.3 cm with a standard deviation equal to 1.8 cm.

(a) If we wanted to determine if the mean head circumference for men was greater than for women, what would the null and alternative hypotheses look like?

(b) Assuming the conditions for a t-test are met (and population variances are equal), calculate the test statistic.

(c) Calculate the p-value.

(d) Make a conclusion based on a 5% significance level and interpret the result in the context of the problem.

Answers

1. (a) Parameter: $\mu_M - \mu_F$, statistic: $\bar{x}_M - \bar{x}_F = 1.1$
- (b) $s_p = 4.26$
- (c) $(-1.053, 3.253)$
- (d) We are 95% confident that the true population difference in mean number of letters printed using the non-dominant hand between males and females is between -.1053 and 3.253 letters. Since 0 is contained in the interval, sufficient evidence does not exist to say that the two mean number of letters differs.

2. (a) $H_0: \mu_M - \mu_F, H_a: \mu_M > \mu_F$
- (b) $t = 4.4, df = 70$
- (c) $P\text{-value} = P(t_{70} > 4.4) = .00002$
- (d) Since .00002 is less than .05, reject the null hypothesis. At the 5% significance level, there is sufficient statistical evidence to indicate that the mean male head circumference is greater than the mean female head circumference.

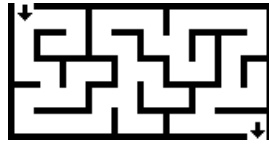
Possible Extensions

1. Demonstrate both the pooled and unpooled two-sample t -test to show students that it's ok to use the unpooled test all of the time, but the pooled test can only be used when the population variances are deemed equal.
2. The assumptions may not be met for a two-sample t -test, so demonstrating the results for a Mann-Whitney U-test may help students better understand the importance of checking assumptions.

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. Maze taken from <http://hereandabove.com/maze/mazeorig.form.html>.
3. Assessment questions from: *Mind on Statistics*, Fourth Edition by Utts/Heckard, 2012. Cengage Learning.
4. Activity background adapted from: <http://en.wikipedia.org/wiki/Maze>.

Activity Sheet
An A-MAZE-ING
Comparison



Background:

A **maze** is a tour puzzle in the form of a complex branching passage through which the solver must find a route. Although the true origins of the maze probably go back to Neolithic times, the earliest mazes we know of were actually architectural monuments built in Egypt and on Crete (an island in the Mediterranean) about 4000 years ago.

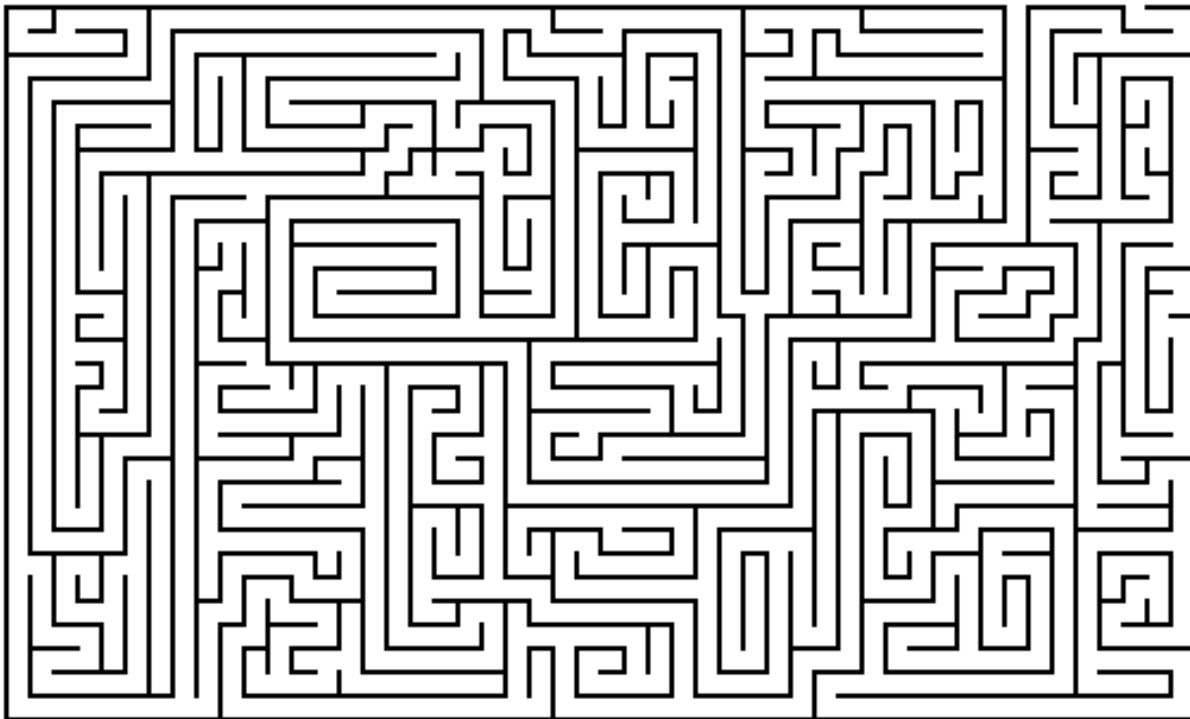
Mazes are often used in psychology experiments to study spatial navigation and learning. We are going to perform a very simplified version of such a psychology experiment.

Problem:

We want to determine if there is a significant difference in the mean number of seconds it takes for males and females to find their way through the following maze.

Instructions:

At an agreed upon time, everyone in the class will be asked to find their way through the following maze.



time taken to complete the maze = _____

gender = M F

Data: Record the class data in the space below.

Male times (in seconds):

Female times (in seconds):

Analysis:

1. Perform a hypothesis test to determine if there is a significant difference in the mean amount of time taken to complete the maze for males and females.

(a) Determine the Hypotheses

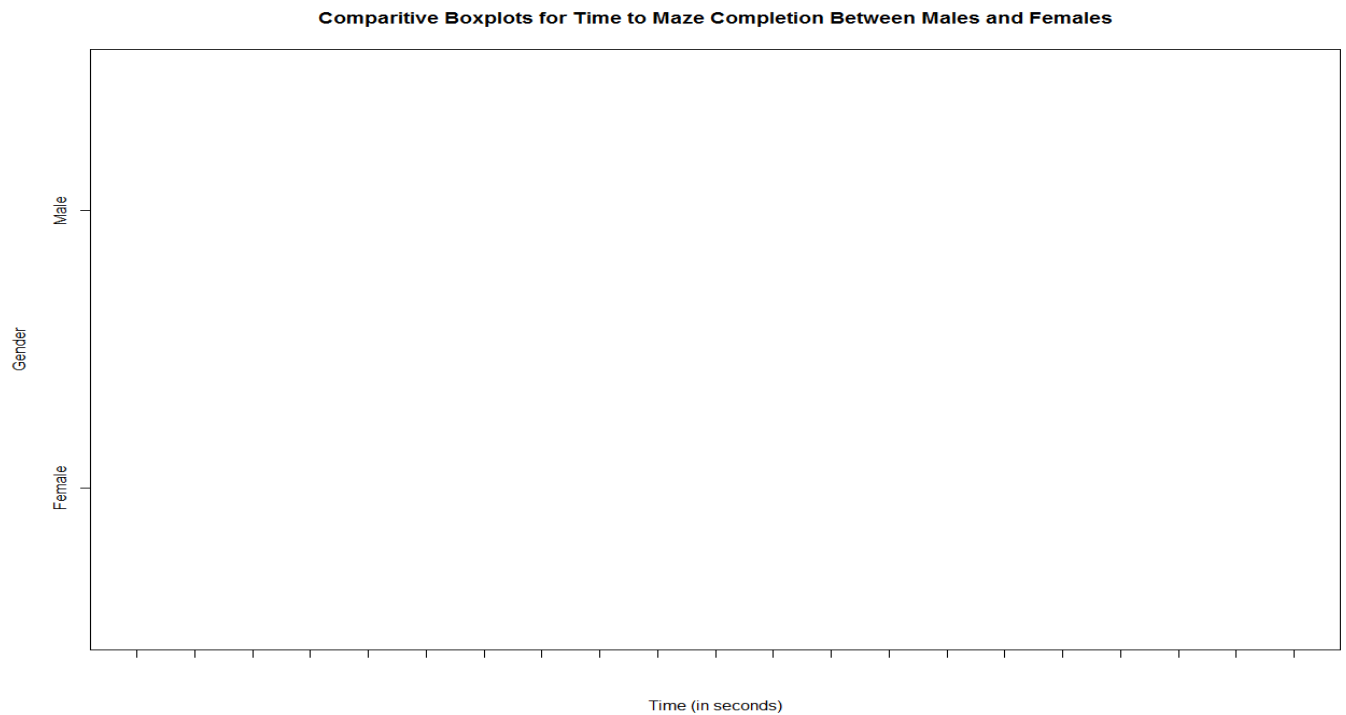
$$H_0 : \qquad \qquad \qquad H_A :$$

(b) Check the Assumptions

(i) Calculate the following descriptive statistics for males and females.

	Males	Females
Sample Size		
Mean		
Standard Deviation		
Minimum		
1st Quartile		
Median		
3rd Quartile		
Maximum		

(ii) Make a side-by-side boxplot for time to maze completion for each group using the descriptive statistics from part (b) above.



(iii) Do the descriptive statistics and comparative boxplots provide evidence for the normality and equality of variances assumptions to be satisfied?

(c) Calculate the Test Statistic

test statistic =

(d) Determine the P-value

p-value =

(e) Provide a Conclusion Based on the P-value and Significance Level

conclusion =

2. Give a practical interpretation of the p-value.

3. Construct a 95% confidence interval for the difference between the male and female mean time to complete the maze. Explain how the confidence interval gives the same conclusion as the hypothesis test performed in Question 1.