What Does the Normal Distribution Sound Like?

Ananda Jayawardhana
Pittsburg State University
ananda@pittstate.edu

Published: June 2013

Overview of Lesson

In this activity, students conduct an investigation to determine the rate of change in the popping of microwave popcorn. Four volunteers will estimate the time in the popping process where the rate of change is at 5 chosen levels. Other students will collect data on number of pops during consecutive 5 second time periods and draw histograms. Students will understand the approximate normal curve for a real data set and the relationship between the rate of popping and the shape of the distribution. This activity also provides an opportunity for students to understand human errors in data collection because not everyone will have the same number of frequencies in the corresponding classes of their histograms.

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 B 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.
6. Attend to precision.
7. Look for and make use of structure.

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. 4. Use the mean and standard deviation of a data set to fit it to a normal distribution and to estimate population percentages. Recognize that there are data sets for which such a procedure is not appropriate. Use calculators, spreadsheets, and tables to estimate areas under the normal curve.

NCTM Principles and Standards for Mathematical Practice
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.

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.

Prerequisites
Students will know the definitions of qualitative and quantitative variables and the different levels of measurement and corresponding graphical techniques used to represent data. Students will have experience in collecting and recording data and in the construction of histograms. Also students will have a basic understanding of the increasing and decreasing rate of change in a function as its input changes.

Learning Targets
Students will be able to determine an appropriate graph to display and interpret data. Students will also be able to understand how to approximate a normal curve to a real data set using graphical techniques. Students will be able to estimate the slope of a smoothed curve of a histogram using class frequencies.

Time Required
One class period.

Material Required
Microwave oven
Three bags of microwave popcorn from the same box
Wall clock
Pencils, graph paper, and a copy of the Activity Sheet

Instructional Lesson Plan

The GAISE Statistical Problem-Solving Procedure

I. Formulate Question(s)
Begin the lesson by asking the students if they know of any real life data distribution which can be approximated by a bell curve. Give some examples like the height of adult men in the U.S. Discuss how rare it is to be extremely short or extremely tall and how common it is to be of average height.

Draw a normal curve on the board and mark the five points as in Figure 1 below.

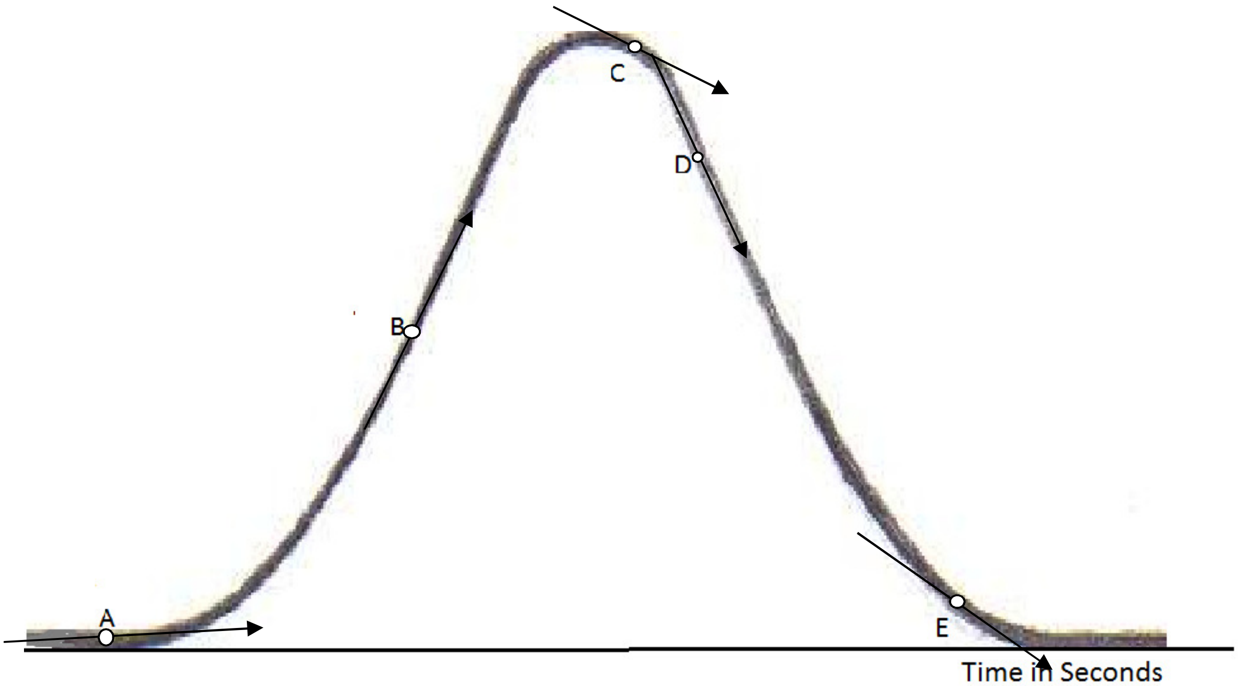


Figure 1. Normal distribution curve.

Ask the students about the slope of the curve at the five different points. Point A is on the curve where the process has just begun and therefore the slope is positive in sign but very small in magnitude. At point B the rate of popping is reaching the maximum. Point C is on the curve just after the peak and the slope is negative with a moderate magnitude. At point D the slope is very close to the maximum in magnitude but negative. Finally, point E is closer to the end of the process but has a slope close to that of point C.

Ask the students about their experience with popping popcorn using a microwave, especially the rate of popping. Ask them if they were to microwave popcorn bags, what variables can be observed and recorded. Ask them what a normal distribution might sound like? Then ask the students to write some questions that they would be interested in investigating regarding popping popcorn. The class generates a set of questions and a volunteer records them on the board.

Some possible questions might be:
1. Will the histogram of the time to pop have a bell shape?
2. How long will it take to hear the first popping sound?
3. How long will it take to pop about 50% of the corn kernels?
4. Will the answer to question (3) be the same for all three bags if they came from the same box?
5. How long will it take to reach the 5 marked points A, B, C, D, and E on the curve?
6. How long will it take to hear the last popping sound?
7. What is the area under the curve?

II. Design and Implement a Plan to Collect the Data

Prepare an Activity Sheet (see pages 11-13) prior to the class. Place the microwave in the classroom so that all students can hear the popping sounds and set the clock to zero before starting the experiment. Ask 4 volunteers to estimate the time that has passed when the popping reaches the 5 marked points on the normal curve and ask the other students to count the popping sounds in 5 second periods and record the results in a table similar to Table 1 below (Table B on the Activity Sheet).

Table 1. Example class results.

Time in Seconds	Tally	Number of Popping Sounds
20-25		0
25-30	*	1
30-35	**	2
35-40	***	3
40-45	*****	5
45-50	***************	15
50-55	******************	18
55-60	******************	18
60-65	********************	20
65-70	******************	18
70-75	***************	15
75-80	***********	11
80-85	**********	10
85-90	********	8
90-95	******	6
95-100	******	6
100-105	*****	5
105-110	**	2
110-115	*	1
115-120	*	1
120-125	*	1

Ask the students the best way to graphically represent the data. Note: If the experiment is conducted for more than 2 minutes students will not be able to taste a normal distribution and safety could be an issue. Also, ask the students whether anyone is allergic to popcorn. As an alternative, the teacher can video the experiment at home and play the video in class. One advantage of the video is the ability to replay and let students check their data for accuracy. Another possibility is to conduct the experiment in class and ask a volunteer to video and load the video to a class computer.

Repeat the experiment with the second and third bags of popcorn and ask the students to record the frequencies of popping in second and third copies of Table B (see pages 14 and 15).

Draw a table as given below on the board with 10 rows. Ask the 4 volunteers who estimated the time to reach each point to write their estimates on the board. Assign 6 other students to complete the extra 6 rows after they analyze the data.

Table 2. Class data collection table.

Volunteer	Time to reach point A	Time to reach point B	Time to reach point C	Time to reach point D	Time to reach point E
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					

III. Analyze the Data

Ask a volunteer to copy his/her frequency table (first Table B on Activity Sheet) on the board or use an Elmo to project an activity sheet. Let the students discuss why different students have different frequencies but the same general frequency pattern appears in the data. Discuss the possible errors in data collection. Facilitate the discussion about the shape of the tallies after rotating the distribution 90 degrees counter clockwise. Ask the students what graphical presentation is the best to represent this data. Ask them about the different types of data and appropriate graphical presentations for them. Ask them to draw histograms for their data in Table B, draw a smooth bell curve over their histogram and mark the corresponding 5 points on the curve as in the diagram given on the activity sheet. Label this graph as Graph 1. Direct them to draw vertical lines through the 5 points and the turning point at the top of the curve to the horizontal axis and estimate the places where these lines cross the horizontal axis as demonstrated in Figure 2 on the next page. Ask them to write above each of the 5 points the sign of the slope and size of the magnitude of the slope as almost level, medium, or steep. Ask them to compare their estimates of the time passed to reach each of the 5 points on the curve. Ask 6 volunteers to complete the last 6 rows of the class data collection table (Table 2) on the board.

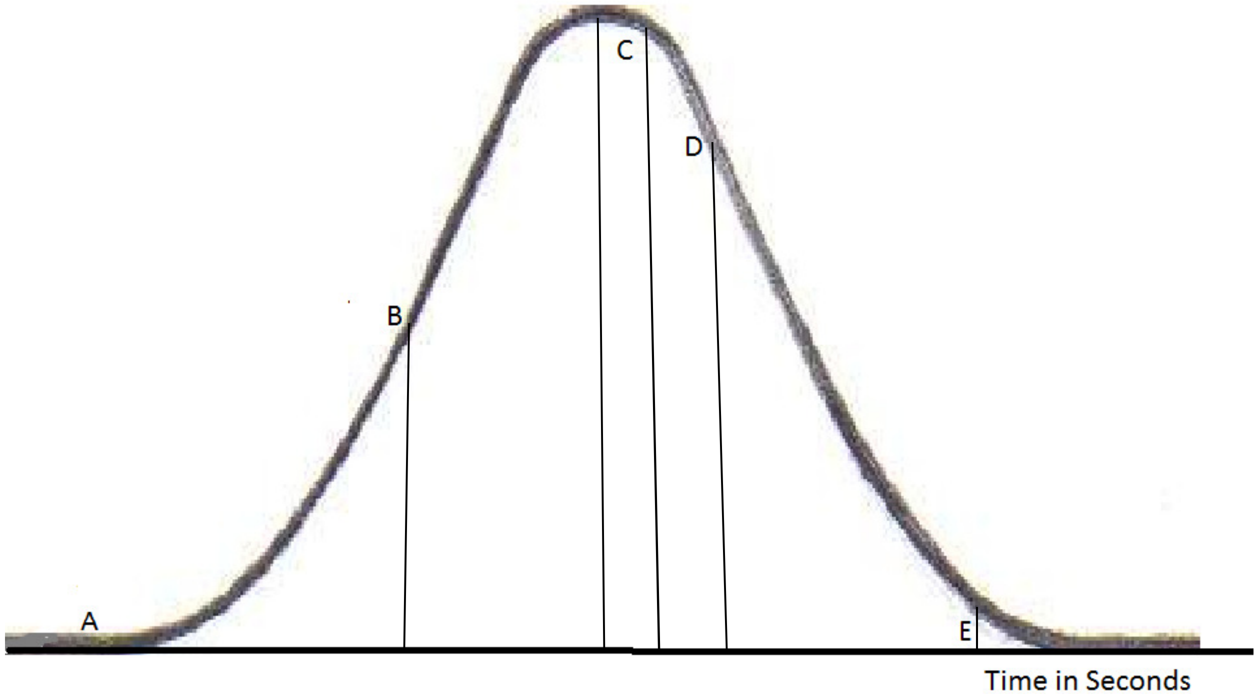


Figure 2. Demonstration of how to draw vertical lines through the five given points and the turning point of the curve

IV. Interpret the Results

1. Will the graph be approximately normal?
   Corn kernels have a certain percentage of water inside them and it is sealed from the outer shell. Microwaves heat up the water and the water turns to steam and the steam pressure breaks the outer shell of the kernel and pops them. The amount of time needed to pop the corn depends on several factors. To name a few, size of the kernel, percentage of water inside the kernel, rupture strength of the shell, and the intensity of the microwaves. In this experiment the only factor we can assume as uniform is the intensity of the microwaves, even though it may be different from one location to another inside the microwave oven. If the kernels were taken from the same cob, one can assume them to be uniform from many aspects. There will still be some natural variability even among the kernels of the same cob. It is likely that commercially packed popcorn will have kernels from different plants from different corn fields. These factors add variability to the popping time. Regardless, there will be an average popping time for most of the kernels. Few will pop quickly and few will take a longer time. More and more will pop closer to the average time producing a bell curve. In the given example the data has a bell shape but it is positively skewed, meaning the longer tail is to the right.

2. How long will it take to hear the first popping sound?
   The answer to this question can change from bag to bag. Most of the students will have the same answer.

3. How long will it take to pop about 50% of the corn kernels?
   Normal distributions are symmetric about the median. The peak of the normal curve corresponds to the median of the variable on the horizontal axis. Ask students to find the class which contains the median and approximate a median time using the data from Table B.

4. Will the answer to question (3) be the same for all three bags if they came from the same box?
   To answer this question three bags should be popped and the results should be recorded in three different frequency tables. Let the students answer the question using the 3 medians calculated using the same classes. Initiate a discussion of natural variability. Ask students to speculate about results using bags from different brand names and from different boxes from the same brand name.

5. How long will it take to reach the 5 marked points on the curve?
   Ask the students to mark the points A, B, C, D, and E on the smoothed curve on their Graph 1. At what times of the popping, does it reach these 5 points? Do they agree with each other's estimates and with the estimates of the 4 volunteers who only estimated the time to reach the 5 points?

6. How long will it take to hear the last popping sound?
   Since the experiment has been stopped (censored) at 2 minutes, it is not possible to estimate the time of the last popping sound. Let the students speculate on the answer.

7. What is the area under the curve?
   Ask the students to find the areas of the rectangles in Graph 1 and add.

Assessment
The following table represents the height distribution of 80 adults in a service club.

Height in Inches	Tally	Frequency
60-62	//	2
62-64	~~////~~ //	7
64-66	~~////~~ ~~////~~ ~~////~~ ///	18
66-68	~~////~~ ~~////~~ ~~////~~ ~~////~~ ~~////~~ /	26
68-70	~~////~~ ~~////~~ ~~////~~ ////	19
70-72	~~////~~ /	6
72-74	/	1
74-76	/	1

1. Draw a histogram to represent the data.

2. Draw a smoothed curve to represent the distribution. Does the curve look like a graph of a normal distribution?

3. What height value will correspond to the point B as in the set of 5 points?

4. Estimate the median.

5. Mark the sign of the slope on the curve for Heights 61, 65, 67, 69, 71, 73, and 75 and give the size of the magnitude of the slope as almost level, medium, or steep.

6. Approximate the area under this curve.

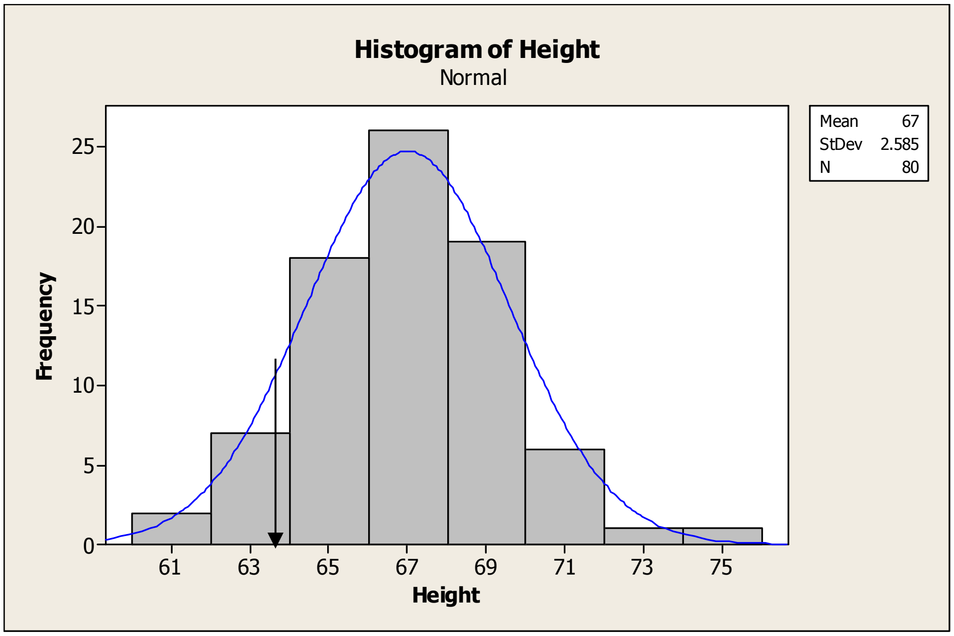
7. Draw the histogram as a percentage on the y-axis and recalculate the area.

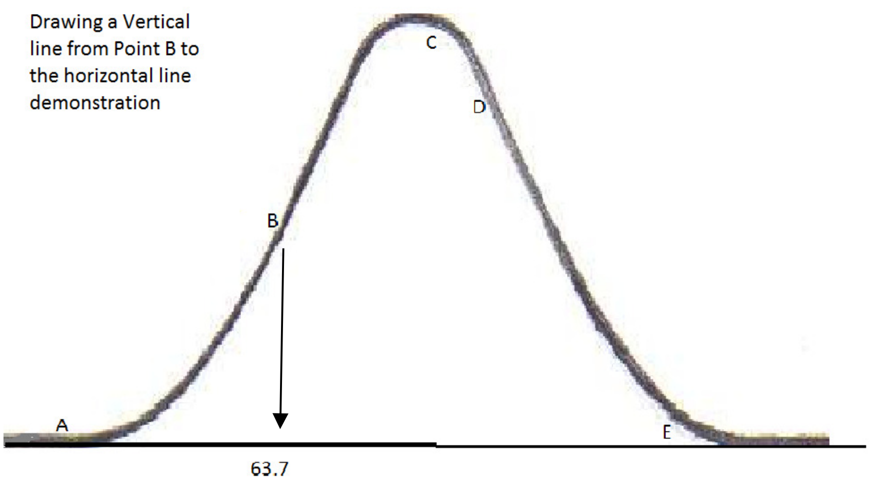
8. Divide the percentages by the total area from part 7 and draw a density histogram.

9. In the density histogram, estimate the slope at Height equals 63. What will be the area under the density curve?

Answers

1.





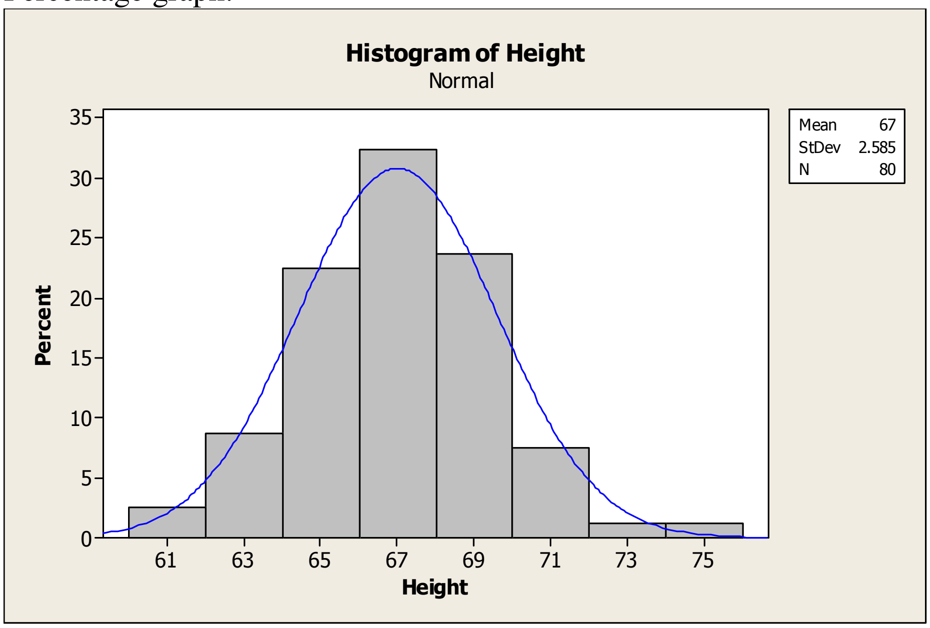
2. The distribution is close to a normal distribution but slightly skewed to the right.
3. An estimate is about 63.7.
4. There are 27 observations before the class 66-68 and 27 observations after the class 66-68. Class 66-68 has 26 observations. If 13 are assigned to the left of the median and 13 are assigned to the right of the median, the two sides of the median have about half the data in each side. If the length of the class 68-66=2 is divided according to the ratio 13:13, then $66+2\times\frac{13}{26}=67$ is a reasonable estimate for the median.

5.

Height	61	65	67	69	71	75
Slope	Positive	Positive	Slope is close to zero	Negative	Negative	Negative
Magnitude of the slope	Medium	Steep	Almost level	Steep	Steep	Almost level

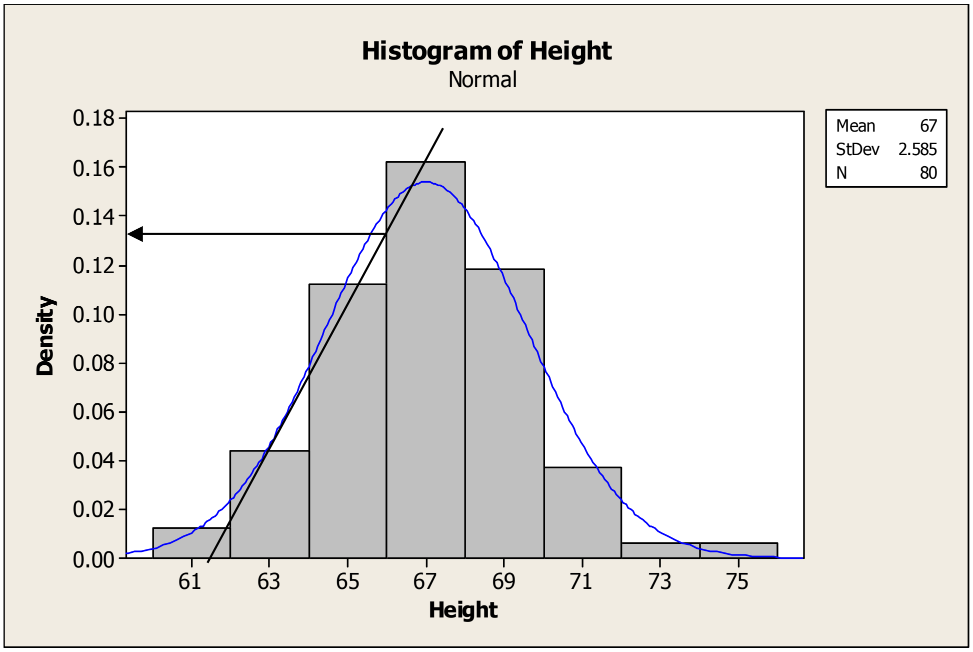
6. Sum of the area of the rectangles:
   Area = $2\times2+2\times7+2\times17+2\times26+2\times19+2\times6+2\times1+2\times1=158$

7. Percentage graph:



Area under the curve = 2

8.



9. Slope $\approx \dfrac{0.135-0.00}{66-61.5} = 0.03$. Area under the density curve = 1.

Possible Extensions

1. This experiment could be used to introduce Z-scores, the standard normal distribution, and how to use the tables of the standard normal distribution to find area under the curve. Questions regarding popping time such as "What percentage of the popcorn will be popped before 40 seconds?" or "How long will it take to pop 75% of the popcorn?" could be answered.
2. Using the mean and standard deviation of the data from this experiment, the teacher can simulate normal random data and demonstrate how to use computer software to simulate the popcorn experiment.

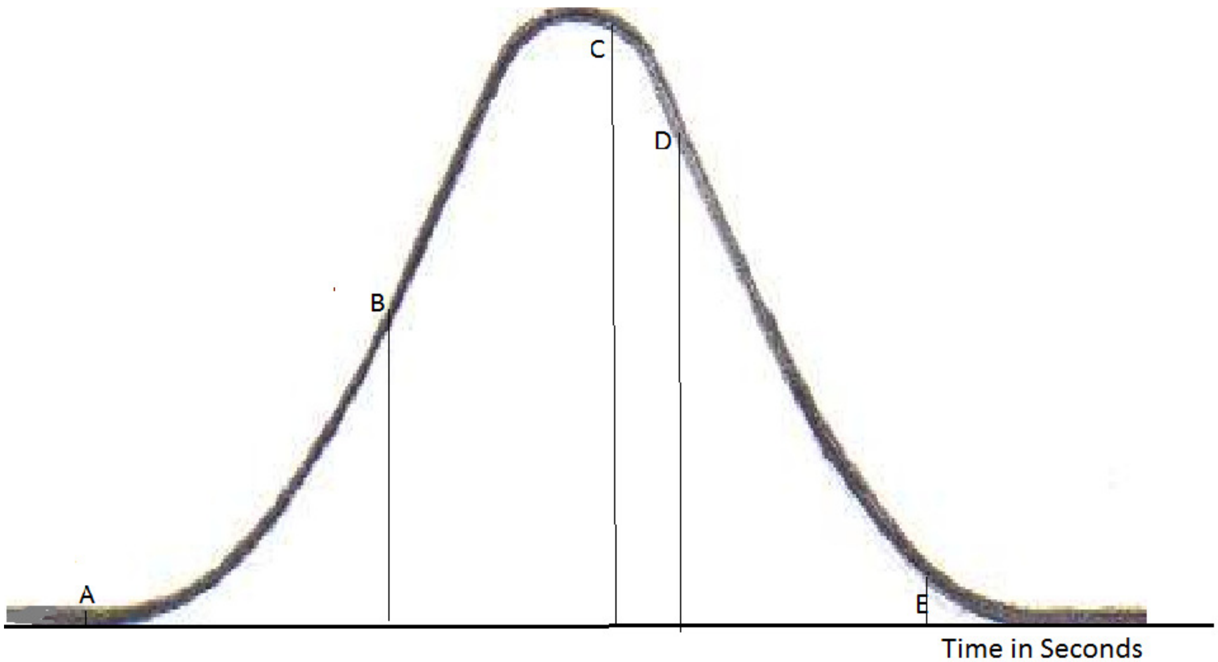
References

Adapted from an activity created for the *American Statistical Association Meeting Within a Meeting Program for Middle School Teachers* (2008).

What Does the Normal Distribution Sound Like? Activity Sheet

Name:……………………………………………………………………..

1. Above each of the 5 points A, B, C, D, and E, write the sign of the slope as positive or negative and the rank of the magnitude of the slope as almost level, medium, or steep.



2. Starting from the left record the time to reach each point if you are one of the volunteers. Tracing on the curve with the sound will help. If it is hard to trace and record, you can ask another student to keep track of the time and record when you give a signal.

Table A

Time to reach point A	Time to reach point B	Time to reach point C	Time to reach point D	Time to reach point E

3. If you are not a volunteer to fill in the table in part (2), count the number of pops in each time interval and record in the following table. If you hear a pop sound exactly on top of a class boundary, add the count to one of the classes in either side.

Table B

Time in Seconds	Tally	Number of Popping Sounds
0-5		
5-10		
10-15		
15-20		
20-25		
25-30		
30-35		
35-40		
40-45		
45-50		
50-55		
55-60		
60-65		
65-70		
70-75		
75-80		
80-85		
85-90		
90-95		
95-100		
100-105		
105-110		
110-115		
115-120		
120-125		

4. Draw a histogram using the same classes as used in Question 3.

5. Draw a smooth curve on top of the histogram. Discuss whether your curve looks like a normal distribution.

6. Estimate the median popping time.

7. Mark the corresponding five points on the curve given in (1) on the curve you drew in part (5).

8. Draw vertical lines through the 5 points to the horizontal axis.

9. Estimate the five corresponding points on the horizontal axis. If you are a volunteer to record these estimates on the board, please do so.

10. Estimate the area under the curve.

Second Copy of Table B

Time in Seconds	Tally	Number of Popping Sounds
0-5		
5-10		
10-15		
15-20		
20-25		
25-30		
30-35		
35-40		
40-45		
45-50		
50-55		
55-60		
60-65		
65-70		
70-75		
75-80		
80-85		
85-90		
90-95		
95-100		
100-105		
105-110		
110-115		
115-120		
120-125		

Third Copy of Table B

Time in Seconds	Tally	Number of Popping Sounds
0-5		
5-10		
10-15		
15-20		
20-25		
25-30		
30-35		
35-40		
40-45		
45-50		
50-55		
55-60		
60-65		
65-70		
70-75		
75-80		
80-85		
85-90		
90-95		
95-100		
100-105		
105-110		
110-115		
115-120		
120-125		