Algebra: Ages of Famous People Lab

Below is a list of famous people and their ages in 2014.

	Estimated Age (years)	Actual Age (years)	Deviation (Estimated – Actual)	In 2014
Michelle Obama				50
Justin Bieber				20
Sofia Vergara				42
Chelsea Clinton				34
Will Smith				46
Adele				24
Ellen DeGeneres				56
Lady Gaga				28
Michael Jordan				51
Oprah Winfrey				60
Jon Bon Jovi				52
Bruce Springsteen				64
Simon Cowell				54

THIS IS THE END OF THIS FILE

Algebra: Ages of Famous People Lab

Below is a list of famous people and their ages in 2014.

	Estimated Age (years)	Actual Age (years)	Deviation (Estimated – Actual)
Michelle Obama			
Justin Bieber			
Sofia Vergara			
Chelsea Clinton			
Will Smith			
Adele			
Ellen DeGeneres			
Lady Gaga			
Michael Jordan			
Oprah Winfrey			
Jon Bon Jovi			
Bruce Springsteen			
Simon Cowell			

THIS IS THE END OF THIS FILE

Algebra 1 Probability

1. **Baseball:** A baseball player's batting average is the ratio of hits to times at bat. The batting average could also be considered as the relative frequency of successes (obtaining a hit) among the batter's official trips to the plate. Batting average can also be interpreted as the probability that the player will get a hit in a randomly selected time at bat. Let's use Ty Cobb's statistics from his more than two decades of play in the major leagues.

- Times at bat 11,429
- Hits 4191
- Singles 3052
- Doubles 724
- Triples 297
- Homeruns 118
- Strikeouts 357

So, Cobb's batting average is $\frac{4191}{11429}$ = .367. In other words the probability that he gets a hit is 36.7%.

a. What is the probability that Cobb would get a single in a random trip to the plate?

b. What is the probability that Cobb would get either a single or a double in a random trip to the plate?

c. What is the probability that Cobb would **not** strike out in a random trip to the plate?

d. Use the net or the library to get similar career statistics for Barry Bonds or Pete Rose. Which player would you say is the better player and explain why.

2. **Fair Coin:** A fair coin is tossed 4 times. What is the probability of getting 2 heads and 2 tails in the four tosses?

a. First let's simulate the experiment by actually tossing coins. Flip a fair coin 4 times and record your data. eg. HTTT Flip the coin 4 more times and record the data. Do this 25 times recording the data for each group of 4.

# of tosses	Data: eg. HTTT
1	
2	
3	
4	
5	
6	

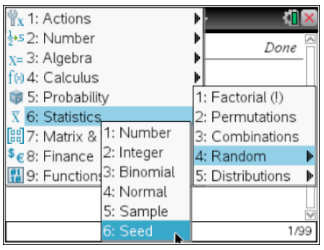
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	

b. Count the number of times you flipped 2H and 2T. Using your data, what is the experimental probability of getting 2H and 2T?

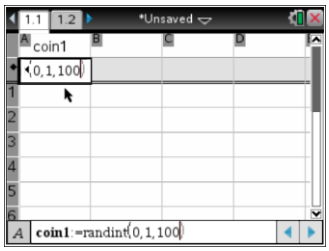
c. Combine your data with data from the others in the class. Now, what is the experimental probability of getting 2H and 2T?

The more tosses we do, the better the approximation will be, but what is the actual probability of this event happening?

Using TI Nspire will allow us to produce as many trials as we want. It is always good to "seed" the random number generator.
In a calculator page go to Menu—Probability—Random--seed

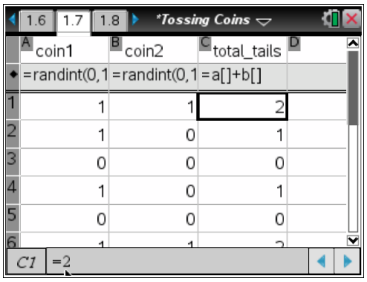


In a spreadsheet we can generate 1 coin toss 100 times using **=randint(0,1,100)** Let 0 = Head and 1 = Tails

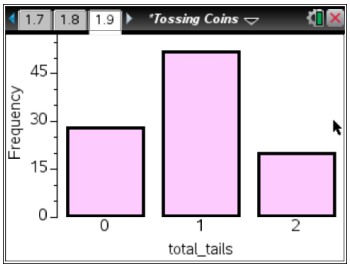


The parameters for randint are the lower bound number, higher bound number, # of integers you would like say (0,1,100). It automatically puts them in the list that was highlighted.

When tossing 2 coins 100 times each, what is the probability of getting T T? note: 2 in the total column indicates T T; a 0 would indicate H H.



A graph will help visualize the data. Insert a data and statistics page. Choose Total Tails for the x- axis. Change the dot plot to a bar graph and use the Force Categorical X and bar chart commands under Menu. Is this the image you had in mind? Go back to the spreadsheet and redo the randint functions. NSpire helps to visualize these experimental probabilities, but can we prove it?



Let us use the 4 coin toss problem to determine the theoretical probability. List the possible patterns of heads and tails that could result from 4 tosses. (Note: this is also called the sample space.)
Eg. HHHH HTHH, etc. How many of these patterns are there?

List the outcomes that contain exactly 2 heads and 2 tails.

Since the coin is fair, each of these possibilities is as likely to occur as any other outcome so the probability is $\frac{\#of\ ways\,to\,get\,2H\,\&\,2T}{total\,\#of\ outcomes}=\frac{6}{16}=\frac{3}{8}$. How close is our experimental solution?

3. Determine the probability of tossing 1 H and 3 Tails. Explain your work.

4. Determine the probability of tossing 3 H and 2 Tails in 5 tosses of a fair coin. Explain your work.

5. If you toss a coin three times the **sample space** is {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT} so the size of the sample space is 8.

a. What was the size of the sample space in problem 2a?

b. What was the size of the sample space in problem 4?

c. If you tossed the coin 6 times, what is the size of the sample space?

d. If you tossed the coin 8 times, what is the size of the sample space?

e. If you tossed the coin 100 times, what is the size of the sample space?

f. If you tossed the coin **n** times, what is the size of the sample space?

Simulating Experiments:

1. **A girl in the Family**: Suppose you want to determine the likelihood of a couple's having a girl among their first 4 children. You may be tempted to say it is very likely or yes, of course, they will have a girl if they have four children. But be careful, a friend of mine, Mike O'Neil has 6 children and they are all boys! So, how likely is it that our mythical couple will have a girl among their first 4 children? We can do an experiment that models the situation and gives us a good estimate for the problem. Since the chances (probability) of a couple having a boy or a girl is approximately equal, we can toss a fair coin to simulate having a baby. (It is much quicker too!) Let heads represent having a girl baby. Now flip the coin until a head appears or until the coin was flipped 4 times, whichever comes first. You and your partners must do this many times to get a good approximation. Record your trials below. Keep track of the number of successes and the number of total trials.

Another way to simulate this problem is to roll a die many times. If the number is odd, it is a girl and if it is even, it is a boy. Can you think of other ways to simulate this experiment?

Using the RandInt function, your TI-NSpire also be used to simulate this experiment.

The parameters for randint are the lower bound number, higher bound number, # of integers you would like say (0,1,4). It automatically puts them in the list that was highlighted.

If you hit the enter button over and over you can simulate the experiment. What is your experimental probability of having a girl among the first 4 children?

2. **Coin Toss Experiment**: Toss a coin 10 times. What is the likelihood of a run of at least 3 consecutive heads or 3 consecutive tails?

a. What are the assumptions in our experiment?

b. Determine a reasonable simulation model and conduct the experiment. According to your simulation, what is the probability of the event?

c. Add your data to data from the other groups. Recalculate the probability. Do you think this is reasonable? Why?

Counting Principles:

1. To meet dress code you have 2 sport coats and 4 pairs of non-jean pants. How many different coat-pants outfits can you make? Explain your work.

2. What if you add 5 dress shirts to the mix in #1? Now how many different outfits can you make?

3. The five of you line up for pizza. How many different ways can you line up?

3b. What if the entire school of 570 students lined up? How many different ways could this be done?

4. How many different license plates can the state of CT make if they decide to use 2 letters followed by 3 digits? Do you think this will be enough plates for all drivers in the state?

4b. Suppose the plates could have either a digit or a letter in each of the 5 spaces. How many different plates could be made?

5. Suppose the letters of Hotchkiss are used to form "words".

a. How many 9 letter "words" can be formed?

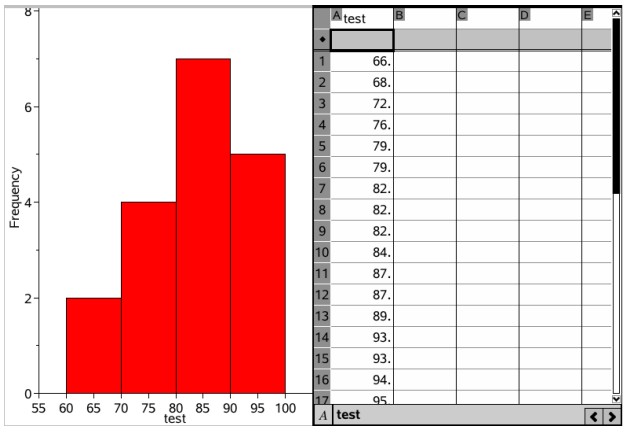
b. How many 8 letter "words" can be formed?

c. How many 5 letter words begin with a vowel and end with a consonant?

THIS IS THE END OF THIS FILE

Applied Calculus Geometric Probability and Integrals Lab based on work in Contemporary Calculus through applications, NCSSM, Janson Publications, 1996.

If the following scores occurred on your last test: 66, 68, 72, 76, 79, 79, 82, 82, 82, 84, 87, 87, 89, 93, 93, 94, 95, and 98. Create a histogram.



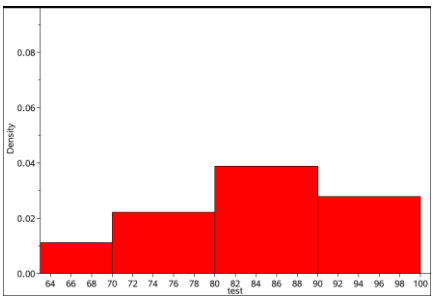
If a student is selected from random from this class, what is the probability that the student's score is between 70 and 90?

But, this can also be thought of as an area problem. If the entire histogram is the sample space, and the rectangles between 70 and 90 constitute the event space.

1. What is the area of the sample space?

2. What is the area of the event space?

Pretty cool but in statistics we use a different vertical scale. We need to rescale the vertical axis so that the total area of the 4 rectangles is 1. To do this divide the height of each rectangle by the total area of all the rectangles, which in this case is 180, the frequency times the width of each rectangle. So the heights of each rectangle would be $\frac{2}{180}, \frac{4}{180}, \frac{7}{180}, \textit{and } \frac{5}{180}$. Change your vertical scale to density and show that the total area under the curve =1.



This is called a **probability histogram** and it makes life much easier since the ratio of $\frac{\text{the area of the event space}}{\textit{the area of the sample space}}$ is simply the area of the event space.

What does the area of each rectangle in the probability histogram represent?

We can approximate the probability histogram by a continuous function called the **probability density function** or **PDF**. There are 2 important characteristics.

1. The function values for the PDF are non-negative so the value of a definite integral of a PDF can be interpreted as an area under the curve.
2. The area under the graph of the PDF over its entire domain is 1.

Suppose you choose a random value for x from the domain of a continuous PDF. Our job is to determine the probability that x lies between 2 values **a** and **b**. In mathematical notation P(a < x < b). We can think of this as an area problem. The region bounded by the graph of f and the x-axis can be thought of as the sample space. So, what do you think is the event space?

If you thought that it is the area under the PDF between x = a and x = b, then bravo! This area is, of course, a definite integral.

$P\ (a < x < b) = \int_a^b f(x)\,dx$ But, how do we determine f(x)?

From http://en.wikipedia.org/wiki/Probability_density_function

Informally, a probability density function can be seen as a "smoothed out" version of a histogram: if one empirically samples enough values of a continuous random variable, producing a histogram depicting relative frequencies of output ranges, then this histogram will resemble the random variable's probability density, assuming that the output ranges are sufficiently narrow.

Any function f that describes the probability density in terms of the input variable x is a probability density function if and only if it is non-negative and the area under the graph is 1:

$$f(x) \geq 0 \forall x \wedge \int_{-\infty}^{\infty} f(x)\,dx = 1$$

The actual probability can then be calculated by taking the integral of the function f(x) by the integration interval of the input variable x.

For example: the probability of the variable X being within the interval [4.3, 7.8] would be

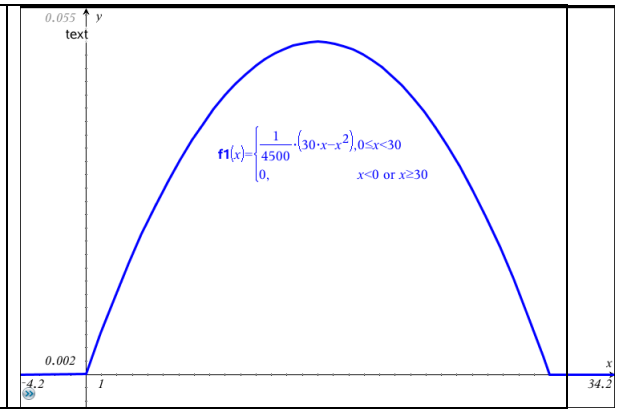
$$\Pr(4.3 \le X \le 7.8) = \int_{4.3}^{7.8} f(x)\,dx.$$

From Contemporary Calculus through applications, page 326

Researchers in a curriculum project have gathered data for the number of minutes required by a sample of students to solve a particular calculus problem. From this data a histogram is constructed, with the # of minutes on the horizontal axis and the # of students requiring that number of minutes on the vertical axis. This histogram is then rescaled to produce the probability histogram. Below is the continuous PDF that approximates the probability histogram.

$$f1(x) := \begin{cases} \frac{1}{4500}\cdot\left(30\cdot x - x^2\right), & 0<x \text{ and } x<30 \\ 0, & x\le 0 \text{ or } x\ge 30 \end{cases}$$

x represents the time it takes a student to solve the problem and that all students can solve the problem within 30 minutes.



Note: Probabilities are represented by areas under the curve.

1. Show that f1(x) is a probability density function. In other words, to show that f1(x) is a PDF you must show that all values of f1 are non-negative and that the total area under the graph of f1 from [0,30] is 1.

2. Find the probability that the time required to solve the problem by randomly selected calculus students is between 12 and 15 minutes.

3. Given the function $f(x) = \begin{cases} kx^2 & 0 \le x \le 10 \\ 0 & elsewhere \end{cases}$

a. Find the value of k for which f will be a probability density function.

b. Find P($4 \le x \le 6$) using the result from part a.

THIS IS THE END OF THIS FILE

The Normal Probability Density Function

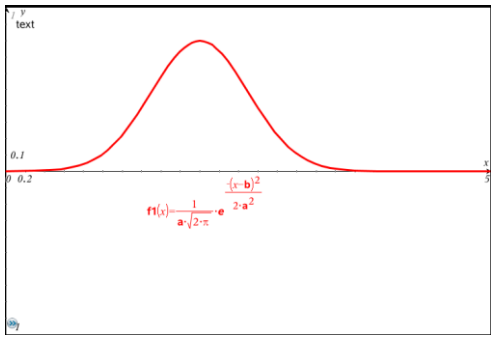
As you learned in the Probability Density Lab, we can approximate the probability histogram by a continuous function called the probability density function or PDF. There are 2 important characteristics.

1. The function values for the PDF are non-negative so the value of a definite integral of a PDF can be interpreted as an area under the curve.
2. The area under the graph of the PDF over its entire domain is 1.

We discovered $f(x) \geq 0 \ \forall x \ and \int_{-\infty}^{\infty} f(x)\,dx = 1$

1. Suppose you found an old clock in the attic. What is the probability that it stopped somewhere between 2:00 and 5:00? Calculus Graphical, Numerical, Algebraic 3rd edition by Finney, et al page 422. Find an appropriate PDF to find your solution.

You are all familiar with the most useful PDF, the Normal PDF sometimes called a "bell shaped curve". Such a graph is at the right.



The Coin Tossing Problem:

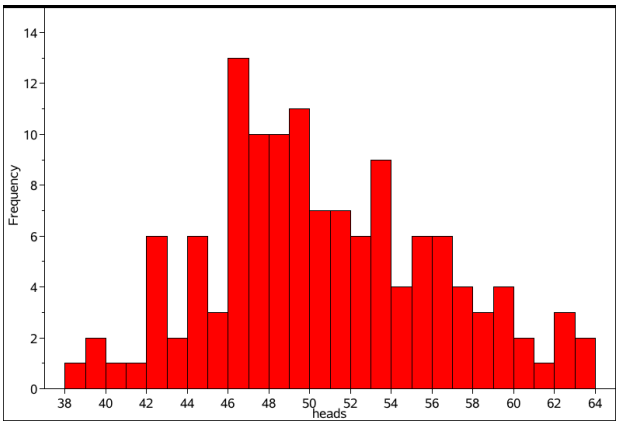
Let's generate some data and look at a histogram. We want to toss 100 coins 100 times and count the number of heads for each toss. In an Nspire spreadsheet, label column A tosses and in cell A1 type the formula =sum(randint(0,1,100)). Note: 0 = tails and 1 = heads If we add them, the total = the number of heads. Why?

Select cell A1 again and now fill down to 100. What does the number in cell A8 represent?

	A tosses	B	C	D	E	F	G	H
1	42.							
2	52.							
3	57.							
4	52.							
5	48.							
6	53.							
7	53.							
8	48.							
9	44.							
10	48.							
11	46.							
12	46.							
13	57.							
14	50.							
15	46.							

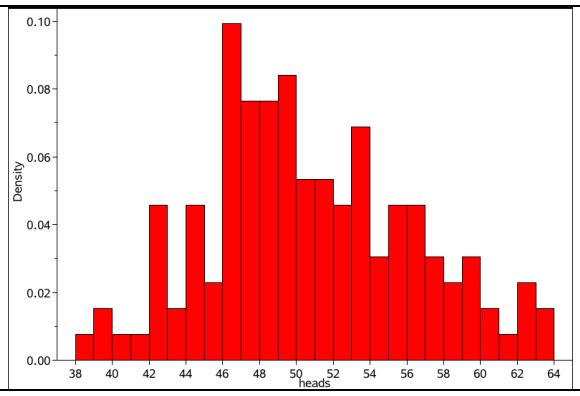
A1 =sum(randint(0,1,100))

Now plot the histogram in a data and statistics window. Show the vertical scale as frequency and set the width of your rectangles (bin width) to 1. Your graph will be different but look something like the one at the right.



1. Based on your histogram, what is the probability that the number of heads in 100 coin flips is
$\leq$ 46? $\leq$50? Is that what you expected? Explain your answer.

Change your histogram and show the vertical scale as density rather than frequency and set the width of your rectangles (bin width) to 1. Do you notice the graphs look identical until you look carefully at the y scale.



2. In a calculator page define f1(x):= e^{-x^2}. Graph it and describe the shape of the graph. Could we use something like this to approximate the PDF above?

The normal probability distribution is the most common of all the continuous probability distributions. The function is:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$ where σ is the standard deviation and μ is the mean of the data.

We want to fit a normal probability distribution model to the histogram. In your calculator window, go to Statistics, Stat Calc and choose 1-var statistics to calculate the mean, μ ($\bar{x}$ on NSpire) and the standard deviation, σ. Define f1(x) using these numbers and the function f(x) above. You will find the standard deviation and the mean under the

VAR button. See f1(x) as defined below. Plot the function on your histogram. Go to Menu, Analyze and then Plot Function.

OneVar *heads*,1: *stat.results*

"Title"	"One−Variable Statistics"
"x̄"	50.26
"Σx"	5026.
"Σx²"	256138.
"sx := sn−1x"	5.97236058097
"σx := σnx"	5.94242374793
"n"	100.
"MinX"	38.
"Q1X"	46.
"MedianX"	50.
"Q3X"	54.
"MaxX"	66.
"SSX := Σ(x−x̄)²"	3531.24

Done

$$f1(x):=\frac{1}{stat.\sigma x\cdot\sqrt{2\cdot\pi}}\cdot e^{\frac{-1}{2}\cdot\left(\frac{x-stat.\bar{x}}{stat.\sigma x}\right)^2}$$

2/99

Now compare your function to the Nspire's normal pdf function. Graph them on the histogram. Impressed? Describe the symmetry of your graph.

If your f1(x) is in fact a PDF, then the area under the curve must = 1. We know all the values in the histogram are positive so that condition is satisfied. Use CAS to find the area under f1(x) from $(-\infty,+\infty)$.

3. Determine the probability of a student being within one standard deviation from the mean. Explain your work.

4. Determine the probability of a student being within two standard deviations from the mean. Explain your work.

5. And finally, determine the probability of a student being within three standard deviations from the mean. Explain your work.

Homework Problems:

1. Goodyear introduces a new passenger car tire. The life span of the tire is normally distributed with an average life of 60,000 miles before wearing out. The standard deviation is 3000 miles. Explain what this means and create a normal pdf function.

a. What is the probability that a randomly chosen tire that will go on your new Chevy Volt will last between 50,000 and 58,000 miles?

b. What is the probability that a randomly chosen tire that will go on your new Chevy Volt will last less than 50,000 miles?

c. What is the probability that a randomly chosen tire that will go on your new Chevy Volt will last more than 70000 miles?

2. Goodyear improves their quality control and the standard deviation drops to 1500 miles and the average life remains 60,000 miles. Describe what this does to your normal PDF function.

a. Now what is the probability that a randomly chosen tire that will go on your new Chevy Volt will last between 50,000 and 58,000 miles?

3. My friend's son is CEO of a customer call center, www.usa800.com/. Assume the mean time for one of their operators to answer a call is 1 ½ minutes with a standard deviation of 30 seconds. Create a normal pdf function to determine the probability:

a. that a call is answered within 1 standard deviation of the mean.

b. that a call is answered in less than 30 seconds.

c. that a call is answered between 45 seconds and 1 minute.

THIS IS THE END OF THIS FILE

Ron Armontrout’s WORD Files ALL – IN - ONE