Sampling Distributions of the Sample Sum and Difference—Dice Rolls

You are going to use Fathom to simulate rolling two dice. For each roll, you will compute the sum and difference of the two dice. Then you'll examine the sampling distributions of the sum and difference.

GENERATE DATA

1. Create a Fathom collection, named Rolls, with 24 cases. Each case will represent a roll of two dice.

Consider randomInteger or randomPick as functions for generating die rolls.

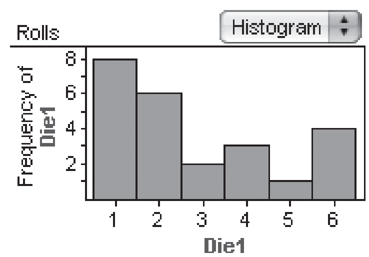
2. Create attributes as shown. Define formulas that randomly roll the dice (have values 1 to 6), and calculate the sum and difference.

Rolls

	Die1	Die2	Sum	Difference
1	6	6	12	0
2	2	5	7	-3
3	6	1	7	5
4	4	2	6	2

Q1 If each face is equally likely on *Die1,* what should the histogram (or dot plot) look like for 24 rolls?

Q2 Make a histogram of *Die1*. Is the distribution approximately uniform, indicating that each outcome is equally likely, or does the distribution follow some other pattern? If it follows some other pattern, why might that be the case?



Q3 Predict what the sampling distribution of the sum and the sampling distribution of the difference will look like for the rolls of two dice. Sketch your predictions.

Q4 Make a histogram of *Sum* and a histogram of *Difference.* Compare your histograms with what you predicted in Q3. Can you describe the sampling distributions in terms of shape, center, and spread based on your 24 rolls of two dice? If so, describe the distributions. If not, explain why not.

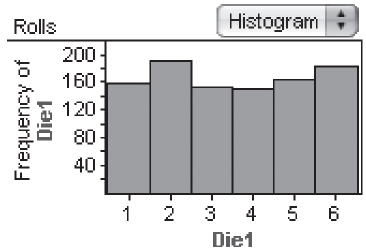
INVESTIGATE

Usually having more cases helps you see trends.

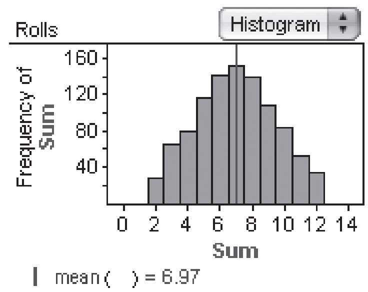
Q5 If each face is equally likely on *Die1,* what should the histogram look like for 1000 rolls?

3. Add cases so that the collection has a total of 1000 cases.

Q6 Does the histogram you plotted for *Die1* look like what you predicted in Q5? Based on this histogram, does it look like each outcome could be equally likely?



Q7 In Q3 you predicted what the sampling distribution of the sum and the sampling distribution of the difference would look like for the rolls of two dice. Compare your predictions with the histogram of *Sum* and the histogram of *Difference*. Are these closer to what you predicted than they were in Q4?



You can choose **Collection | Rerandomize** to look at new samples of 1000 rolls.

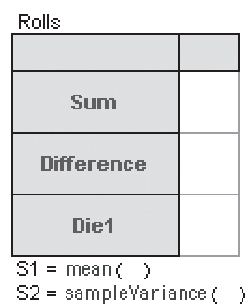
Q8 Describe the sampling distributions of the sum and difference in terms of shape, center, and spread based on your 1000 rolls of two dice.

Q9 How do the sampling distributions of the sum and difference compare with the distribution of a roll of a single die?

Q10 Plot the value of the mean for each distribution. How do the means of the distributions of sums and differences compare with the mean of the distribution of a single roll of a die?

Fathom has the built-in function `variance()`.

4. Use summary tables to compute the variance of each distribution.



Q11 How do the variances of the distributions of sums and differences compare with the variance of the distribution of a single roll of a die? How do they compare with each other? Can you find formulas that relate them?

EXPLORE MORE

1. Investigate the sampling distribution of the sum of three dice rolls and of four dice rolls. Describe and explain your results.
2. Investigate the sampling distributions of the sum and difference for two dice with a number of sides other than 6. Describe and explain your results.
3. Make a scatter plot of *Difference* versus *Sum*. Explain your results.

Consider randomNormal instead of randomInteger to solve this problem.

4. Use what you have learned to answer this question by setting up a similar simulation: Bottle caps are manufactured so that their inside diameters have a distribution that is approximately normal with mean 36 mm and standard deviation 1 mm. The distribution of the outside diameters of the bottles is approximately normal with mean 35 mm and standard deviation 1.2 mm. If a bottle cap and a bottle are selected at random (and independently!), what is the probability that the cap will fit on the bottle?

5. Adjust your simulation in Explore More 4 to answer this question: The distribution of the outside diameters of a set of bottles is approximately normal with mean 35 mm and standard deviation 0.6 mm. The mean of the bottle caps is 36 mm. How small of an error (SD) do you need in your cap-making machine to guarantee that your machine makes a cap that fits 95% of the bottles?

Sampling Distributions of the Sample Sum and Difference—Dice Rolls

Activity Notes

Objectives

- Understanding the concept of sampling distributions of the sum and difference
- Discovering the properties of shape, mean, and *variance* of the sampling distributions of the sum and difference
- Seeing that the mean of the sampling distribution of the sum is approximately the sum of the two population means, and the mean of the sampling distribution of the difference is the difference of the means
- Recognizing that the variance of either sampling distribution is the sum of the population variances

Activity Time: 20–40 minutes

Setting: Paired/Individual Activity (build simulation) or Whole-Class Presentation (use **SumDifference.ftm**)

Optional Document: BottleFit.ftm (Explore More 4 and 5 solutions)

Materials

- *Optional:* One (or two) die for each group of three students

Statistics Prerequisites

- Familiarity with sampling distributions
- Some familiarity with the equally likely outcomes
- Comparing distributions graphically
- Measures of center and spread

Statistics Skills

- Sampling distributions of the sample sum and difference
- Properties of the shape, center, and spread of the sampling distribution of the sample sum and difference
- Central Limit Theorem

AP Course Topic Outline: Part III B, C, D (4, 5, 6)

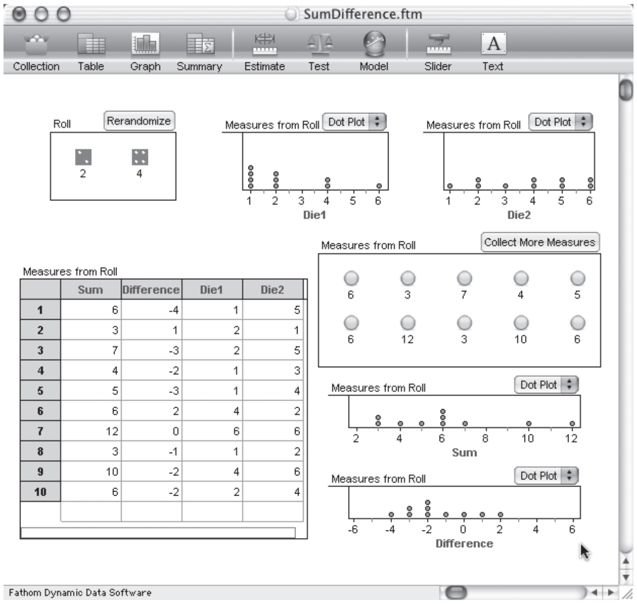
Fathom Prerequisites: Students should be able to make collections and graphs, plot values, find statistics in a summary table, define attributes, and use randomInteger or randomPick functions.

Fathom Skills: Students use formulas to create a collection that represents a random sample, use Fathom to do simulations, and define measures to create the sampling distributions of the sum and difference.

General Notes: With Fathom, students can accomplish the simulation in a very straightforward manner without having to collate classroom data. Once the simulation has been constructed, it can be rerun instantly by rerandomizing the collection. Furthermore, the sample size can be very large (1000) so that the shape of the distributions is very apparent.

Procedure: Some students must actually "see" the sampling done to get the idea before they can use Fathom. If so, you can begin this activity by having them do steps 1–4 by hand. Divide your class into groups of three. Each group will roll their die twice. One person will record the sum of the numbers on the two rolls. Another partner will record the difference. The third person will record the number on each die. For example, if the first roll is 2 and the second roll is 4, the first partner records 6, the second records −2, and the third records both the 2 and the 4. Have them continue to roll until their group has recorded 24 sums and 24 differences. Have them answer Q1–Q4. After that, they do the same thing using Fathom starting at the beginning of the activity again.

If you'd rather students didn't roll the die, you can use the presentation document **SumDifference.ftm** to do the above as a class, or have them use this file on their own. The file is set up differently than the activity. Here there are two dice in the Rolls collection. When you click the **Rerandomize** button, Fathom "rolls" the two dice. When you click the **Collect More Measures** button, Fathom "rolls" the two dice 10 times. The animation shows each roll and updates each dot plot. The values for the icons in the Measures from Roll collection are the sums of the two dice.



To follow the beginning of the activity, double-click the Measures from Roll collection to show its inspector and change the number of measures to 24.

If your students don't need to "see" the hands-on method, they can start right off with the activity.

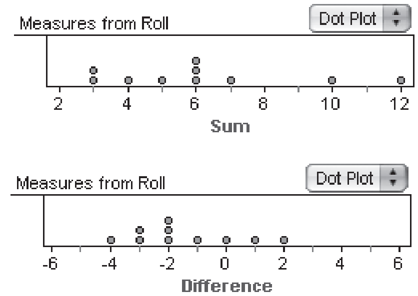
GENERATE DATA

2. Either randomPick(1,2,3,4,5,6) or randomInteger(1,6) will generate die rolls.

Q1–Q2 If each face is equally likely, the histogram for *Die1* should be uniform from 1 to 6, with each bin of height 4. More than likely, none of the histograms will look like students' predictions, because of the small sample size.

Q3 Draw attention to the importance of the prediction in this question. It isn't as important that students predict correctly as it is that they think about what is going on before they make it happen.

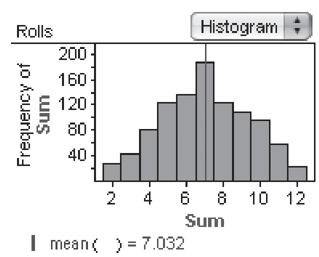
Q4 The histograms will probably not show any pattern, again because the sample is much too small. Looking at these two distributions, students should see that the difference has a hint of triangularity but the sum definitely does not.

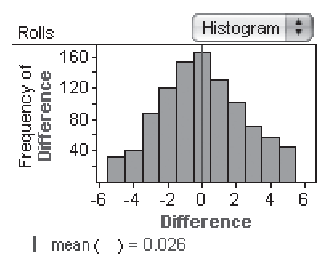


INVESTIGATE

Q5–Q6 If each face is equally likely, the histogram for *Die1* should be uniform from 1 to 6, with each bin height approximately 166.7. More than likely, none of the histograms will look exactly like students' predictions, but they will come fairly close.

Q7–Q8 Both distributions will look approximately triangular. The mean of the sum should be about $3.5 + 3.5 = 7$, or $\mu_1 + \mu_2$. The mean of the difference should be about $3.5 - 3.5 = 0$, or $\mu_1 - \mu_2$. Almost all students will be surprised that the spreads are equal (about 2.4).





Be sure that students rerandomize the collection at least a few times. This helps them see which features of the histograms are real and which are fluctuations.

(*Note:* The standard deviation of the sums or differences is larger than the standard deviation of an individual roll, but the standard deviations do not add. It is the variances that add.)

Q9 The shape of both distributions is triangular while the shape of the population of outcomes is rectangular.

Q10 The sum of the means of the sum and difference should be approximately equal to the mean of the roll of a single die.

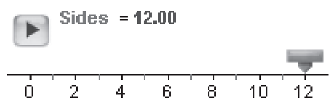
Q11 Both variances should be close to 5.8, or $\sigma_1^2 + \sigma_2^2$. The variance of a single roll should be close to 2.917.

DISCUSSION QUESTIONS

- How is it that the means of the sum and difference distributions are different but the variances are the same?
- How are the means and variances for the sum and difference distributions related to the mean and variance of the distribution of single-die rolls?
- Is there any difference between rolling one die twice and rolling a pair of dice once?

EXPLORE MORE

2. Students could use a slider to redo this simulation for a die with a different number of sides.



Rolls

	Die1	Die2	Sum	Difference
=	randomInteger(1, Sides)	randomInteger(1, Sides)	Die1 + Die2	Die1 – Die2

3. The result is a geometric display of the possible sum and difference combinations. By creating a breakdown plot (hold down the Shift key when dropping the attributes to make them categorical), you can get a sense for the probability of getting any particular sum-difference pair.

4.–5. See the document **BottleFit.ftm** for a solution. For Explore More 4, the probability is 0.739. For Explore More 5, the SE must be at most 0.15.