

Background

In this activity you will be simulating the “mating” of mice (using dice) to investigate the probability parents’ offspring will possess specific traits (i.e. fur color). Some relevant scientific terms:



- DNA (deoxyribonucleic acid) encodes inherited traits on a molecule.
- Genes are segments of DNA that contain instructions for particular traits.
- An allele is a version of a gene that may lead to a trait being expressed.
- Genotype is the two-allele combination in an organism.
- Phenotype is the trait that is expressed (the physical appearance) in an organism.
 - Dominant alleles are always expressed when present.
 - Recessive alleles are not expressed when the dominant allele is present.
- Heterozygous means that the alleles in a genotype are different. Likewise homozygous means that the alleles in a genotype are the same.

Simulation Setup

We will be looking at allele’s related that determine each mouse’s fur color. For this example, mice will either be brown or gray:

- Use capital “F” to represent a dominant allele. When a dominant allele present, it will be expressed and the mouse will have brown fur (“brown” is an example of a phenotype).
- Use lowercase “f” to represent a recessive allele. When only recessive alleles are present, the mouse will have grey fur.

Each parent contributes one of their two alleles at random. As we begin, we’ll make two assumptions:

- Both parents have both a dominant and a recessive allele.
- Each parent is equally likely to contribute either the dominant or the recessive allele.

For example, if both parent mice contribute the dominate allele, the genotype of the child mouse will be “FF”. Fill in the table below to reflect these assumptions.

	Genotype	Phenotype	Chance Parent Contributes “F”	Chance Parent Contributes “f”
Mother				
Father				

To simulate the process of creating a child mouse, we will be using two dice (each of the dice represents one of the parent mice). Answer the questions below as you consider how to do this in a way that reflects the assumptions above.

1. You’ve been given two dice of different colors. What role should the colors play in the simulation?
2. The dice are fair 6-sided dice with the sides numbered from 1 to 6.
 - a. What is the chance of getting any particular value on a given die?
 - b. How do you use that knowledge to assign allele’s according to the assumption stated above?
 - c. You will roll the dice to simulate genotypes for child mice. What other (process oriented) things do you need to consider before beginning the simulation?

Simulation and Empirical Probability

“Mate” your mice 50 times using the dice according to the rules that you’ve set up for the simulation. Record your results for both genotype and phenotype using tally marks in the tables below. Note that “Ff” genotype does not differentiate which allele came from which parent. After you are finished, calculate the frequency and relative frequency of each occurrence. For ease of comparison later, you might use decimals rounded to 4 decimal places for the relative frequencies (it is also acceptable to use fractions).

Genotype

	Tallies	Frequency	Relative Frequency
FF			
Ff			
ff			

Phenotype

Phenotype	Tallies	Frequency	Relative Frequency
Brown			
Grey			

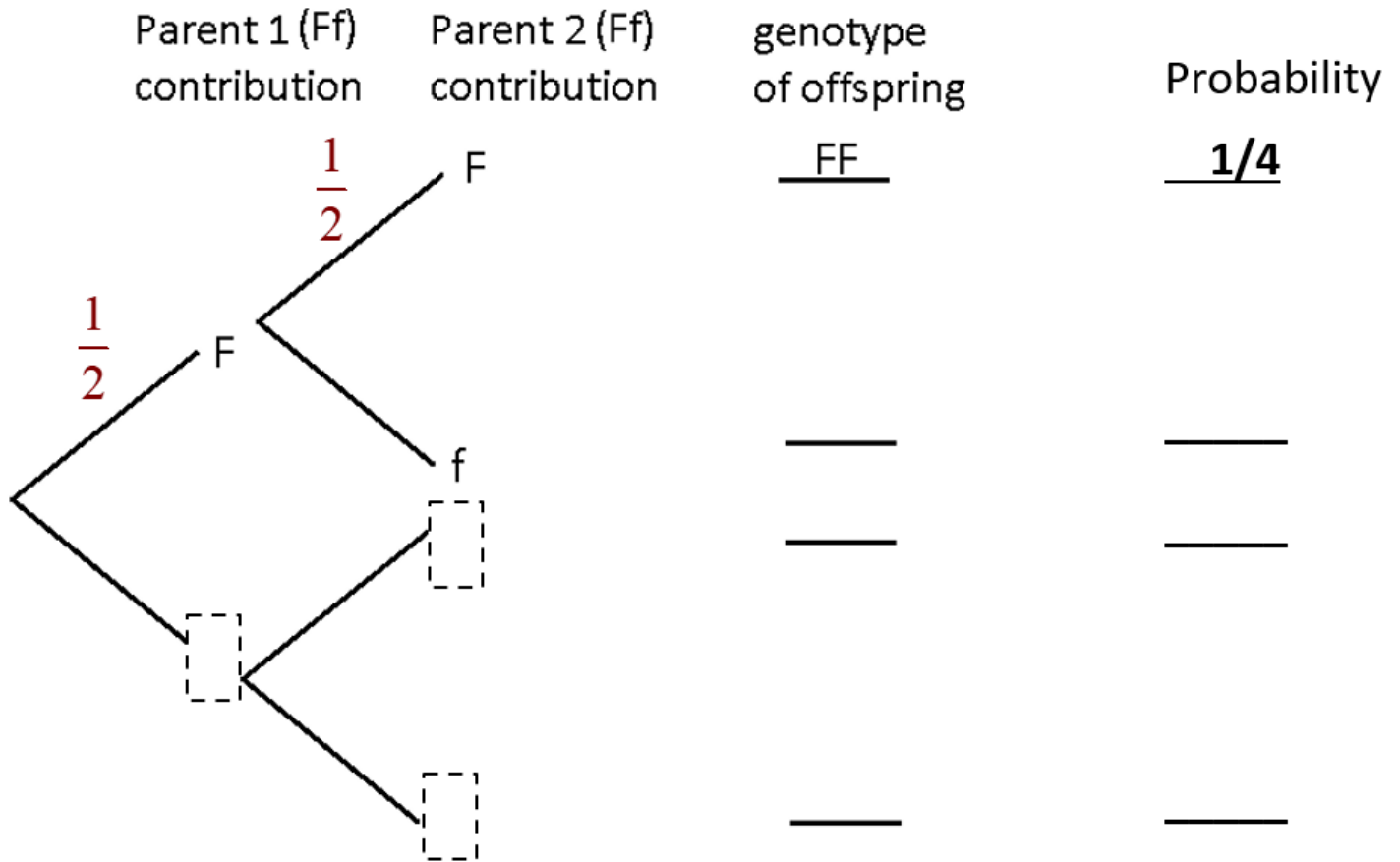
Notes on Statistical Terminology

- **Probability** is a number (fraction or decimal) between 0 and 1 which reflects the likelihood of an event (for example, the **theoretical probability** of an odd die roll is 0.50, or 1/2, since 3 of the 6 equally likely possibilities result in “odd”).
- **Frequency** is simply the number of times each outcome occurred.
- The **relative frequency** is computed by dividing the frequency by the sample size. Relative frequency effectively gives us the **empirical probability** (i.e. a probability estimated based on collected data).
- The **theoretical probability** of an event is the proportion the empirical probability would approach as the number of trials of a probability experiment gets very large.
- A **probability distribution** is a list of all possible outcomes for an experiment (called the **sample space**) along with their corresponding probabilities. (This could be empirical or theoretical)
- **Law of Large Numbers:** Larger sample sizes improve the precision of empirical probability estimates. That is to say, provided an estimate is unbiased, the larger the sample size the closer the estimated (empirical) probability is likely to be to the theoretical probability.

For this simulation it is possible to calculate the theoretical probabilities for the outcomes in the tables above, using a **tree diagram** (shown below). **Tree diagrams are another useful picture and are most useful when outcomes are obtained by taking two or more consecutive “actions” that combine to form an outcome.**

The numbers are very simple here – what you should focus on is the process. Complete the figure below by:

1. Placing the appropriate allele (“F” or “f”) on each remaining **node**, as appropriate.
2. Adding probabilities to all **branches** according to the chance of getting each allele.
3. Listing the genotypes associated with each branching.
4. Multiplying across branches to get the probability for each genotype. Important: When multiplying I recommend NOT reducing fractions. Keeping denominators the same will be helpful when applying the probability distributions to answer other questions.



Based on your theoretical calculations above, complete the probability distributions for genotype and phenotype in the tables below. Then answer the questions that follow.

Genotype	Probability
FF	1/4
Ff	
ff	

Phenotype	Probability
Brown	
Grey	

1. What do you notice about the sum of the probabilities in the probability distribution? Will this always be true? Fully explain why.
2. Do your theoretical probabilities match your empirical probabilities that you got from simulating the experiment? Should they? Why or why not?
3. What could be done to improve upon the “closeness” of the empirical and theoretical probabilities? Explain.

Extension 1

Keep everything about the scenario the say, but change the 50% chance that each allele will be passed to instead have a 40% chance that the dominant allele is passed to the baby, and a 60% chance that the recessive allele is passed to the baby. You can do this with a six-sided dice in the following way:

- Rolls of 1 or 2 are the Dominant allele
- Rolls of 3, 4, or 5 are the Recessive allele
- Rolls of 6 are rerolled until they become a different number.

To save time, you could do 20 rolls instead of 50 if you want. Keep in mind how this impacts estimates via the Law of Large Numbers!

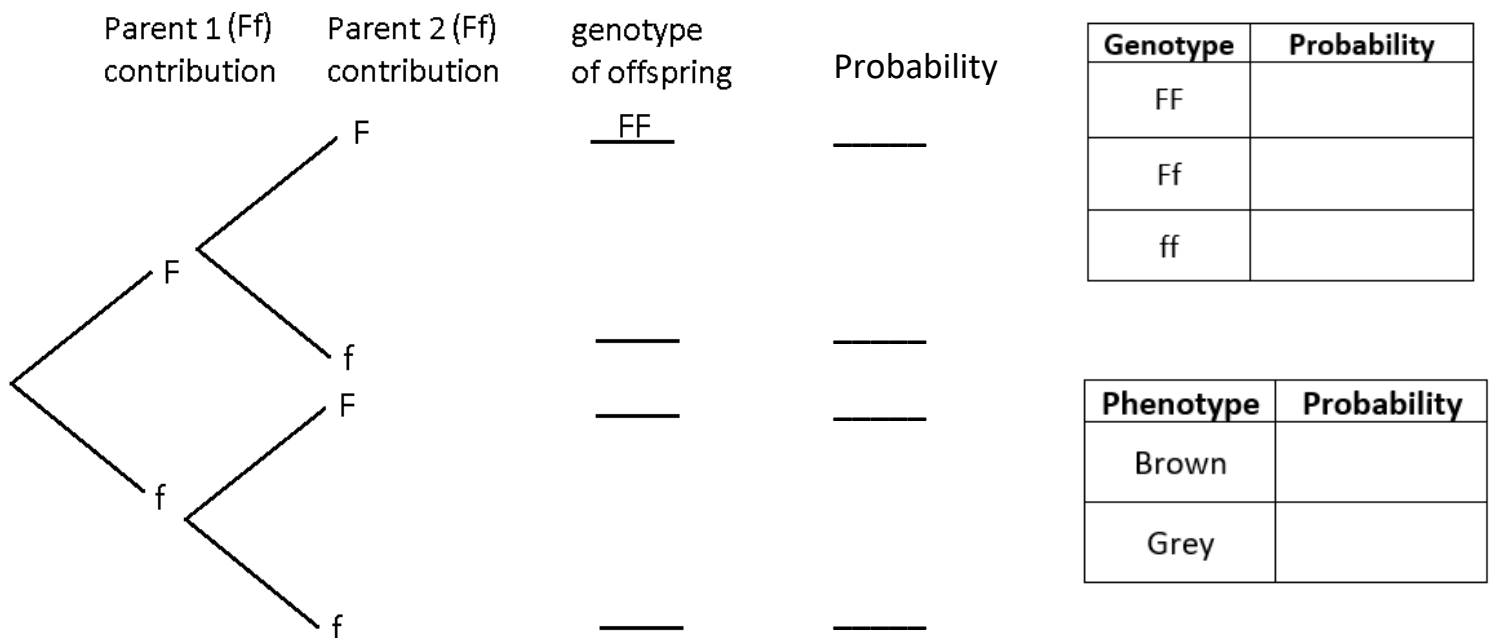
Genotype

	Tallies	Frequency	Relative Frequency
FF			
Ff			
ff			

Phenotype

Phenotype	Tallies	Frequency	Relative Frequency
Brown			
Grey			

Theoretical



Extension 2

Again, keep the original with one change. This time, suppose that the mother has a $\frac{2}{3}$ chance of passing the dominant allele, and that the father has a $\frac{3}{4}$ chance of passing the recessive allele. Again, you can do this with six-sided dice as follows:

- Mother: 1,2,3,4 = Dominant // 5,6 = Recessive
- Father: 1 = Dominant // 2,3,4 = Recessive // 5,6 = reroll

To save time, you could again do 20 rolls instead of 50 if you want. Keep in mind how this impacts estimates via the Law of Large Numbers!

Genotype

	Tallies	Frequency	Relative Frequency
FF			
Ff			
ff			

Phenotype

Phenotype	Tallies	Frequency	Relative Frequency
Brown			
Grey			

Theoretical

