Example 1-1: Random Babies
Suppose that four mothers give birth to baby boys at the same hospital on the same evening. As a very, very sick joke (do not try this at home!), the hospital staff decides to return the babies to the mothers at random! How likely is it (i.e., what is the probability) that at least one mother receives the correct baby?

a) Make a guess for this probability.

This is a classic probability problem known as the matching problem. If you prefer to think of a different context, suppose that four college students bump into each other and drop their cell phones; then each student picks up one of the four cell phones at random.

We’ll take two approaches to answering this question (and many probability questions throughout this course):
1) An approximate analysis via simulation
2) An exact analysis via enumeration

- A process is random if individual outcomes are uncertain but there is a regular distribution of outcomes in a large number of repetitions.
- The probability of any outcome in a random process can be interpreted as the proportion (relative frequency) of times that the outcome would occur if the random process were repeated for a very large number of repetitions.
- A probability can be approximated by simulating (artificially re-creating) the random process a large number of times and calculating the relative frequency of occurrences.

b) Describe how you could use four index cards to simulate this random process and approximate the probability that at least one mother gets the right baby.

c) Use four index cards to conduct 5 repetitions of this random process. For each repetition, record the number of mothers who receive the correct baby.

d) Based on your simulation analysis, report the approximate probability that at least one mother receives the correct baby.
e) Suggest how to determine a more accurate approximation for this probability.

f) Now combine your simulation results with those of your classmates. Record the number and proportion of occurrences for each possible number of matches:

<table>
<thead>
<tr>
<th># of matches</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Count</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

g) Use these simulation results to approximate the probability that at least one mother gets the right baby. (Describe two ways to determine this.)

h) Now we turn to technology in order to perform this simulation more quickly and efficiently. We will use an applet found at: www.rossmanchance.com/applets/randomBabies/Babies.html. (Be forewarned that this applet contains rather graphic images that reveal where babies come from!) Conduct 1000 and then 10,000 repetitions of this simulation. Report the results:

<table>
<thead>
<tr>
<th># of matches</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Count</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

i) Based on the simulation with 10,000 repetitions, report the approximate probability that at least one mother gets the right baby.

- Conducting more repetitions in a simulation analysis generally produces more accurate approximations of probabilities.
- A (rough) rule-of-thumb is that the approximate probability will likely fall within $\pm \frac{1}{\sqrt{N}}$ of the actual probability, where $N$ represents the number of repetitions.

j) Calculate and interpret the value of $\frac{1}{\sqrt{N}}$ for your applet simulation.

k) Click on the bar in the applet output corresponding to 0 matches. Describe what the resulting graph reveals.
Example 1-2: Random Babies (cont.)
Can we use a more mathematical analysis to calculate this probability exactly? Sure! An exact theoretical analysis of this random process would consider all of the possible ways to distribute the four babies to the four mothers. All of the possibilities are listed here:

<table>
<thead>
<tr>
<th></th>
<th>1234</th>
<th>1243</th>
<th>1324</th>
<th>1342</th>
<th>1423</th>
<th>1432</th>
</tr>
</thead>
<tbody>
<tr>
<td>2134</td>
<td>2143</td>
<td>2314</td>
<td>2341</td>
<td>2413</td>
<td>2431</td>
<td></td>
</tr>
<tr>
<td>3124</td>
<td>3142</td>
<td>3214</td>
<td>3241</td>
<td>3412</td>
<td>3421</td>
<td></td>
</tr>
<tr>
<td>4123</td>
<td>4132</td>
<td>4213</td>
<td>4231</td>
<td>4312</td>
<td>4321</td>
<td></td>
</tr>
</tbody>
</table>

a) How many possibilities are there for returning the four babies to their mothers?

b) For each of these possibilities, indicate how many mothers get the correct baby.

c) Count how many ways there are to get 0 matches, 1 match, and so on. Record these in the middle row of the table below:

<table>
<thead>
<tr>
<th># of matches</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td># possibilities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>probability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

d) Determine the (exact theoretical) probability of each event by dividing these counts by your answer to (a). Record these in the last row of the table above.

- The sample space (S) of a random process is a set consisting of all possible outcomes.
- An event is a subset of the sample space (often denoted by a capital letter).
- If the outcomes are equally likely, then the probability of the event is the number of outcomes in the event divided by the total number of outcomes in the sample space.
  o This is sometimes called the classical approach to probability.

e) Determine the exact probability that at least one mother gets the correct baby? Describe two different ways to calculate this. Also interpret this probability value.

f) What is the probability that exactly 3 mothers get the right baby? Explain what this means and why this makes sense.
Example 1-3: Ice Cream Prices
Suppose that you have only 50 cents in your pocket and you want to buy an ice cream cone. The owner of the ice cream shop offers a random price determined as follows: You roll a pair of fair, six-sided dice, and the price is the larger number followed by the smaller number (in cents). What is the probability that you’ll be able to afford the ice cream cone?

Let’s use R to simulate this random process and determine an approximate probability that you’ll be able to afford the ice cream cone, using 10,000 repetitions:

```
N = 10000
d1 = sample(1:6,N, replace=TRUE)
d2 = sample(1:6,N, replace=TRUE)
price = 10*pmax(d1,d2) + pmin(d1,d2)
afford=(price<=50)
sum(afford)/N
```

a) Describe what each line of this small program does.

b) Run this code, and then report the approximate probability along with a rough estimate of its accuracy.

Now we will determine the exact probabilities involved with this random process.

c) List the 36 outcomes in the sample space for rolling a pair of fair, six-sided dice.

d) What is the probability that the ice cream cone costs 33 cents?
e) What is the probability that the ice cream cone costs 32 cents?

f) What is the probability that the ice cream cone costs 23 cents?

g) Circle the outcomes that comprise the event that you can afford the ice cream cone. (Remember that you only have 50 cents.) Then determine the (exact) probability that you can afford the ice cream cone.

h) Is the exact probability close to our R simulation approximation?

i) Suppose that I offer to buy your ice cream cone if the price turns out to be an odd number. What is the probability that I buy your ice cream cone? Is this more likely than not?

Example 1-4: Random Babies (cont.)
Reconsider the random process of distributing babies to their mothers at random, but now suppose that there were 8 babies instead of 4.

a) Make a guess for the probability that at least one mother would get the correct baby. Is your guess larger, smaller, or about the same as the probability with 4 babies/mothers?

b) Do you think we’ll use simulation or enumeration to analyze this random process? Explain why.

c) Examine the following R code (also available from our course webpage). Identify each of the following as a scalar or vector. Indicate how long each vector is.

```r
n
N
x
y
i
nummatches
event
```
c) Run the code with 100,000 repetitions for 4 babies/mothers. Report the approximate probability that at least one mother gets the right baby. Is this within $1/\sqrt{N}$ of the exact probability?

d) Run the code with 100,000 repetitions for 8 babies/mothers. Report the approximate probability that at least one mother gets the right baby. Also give an estimate of the uncertainty with this approximation. How does this probability compare to the probability with 4 babies/mothers?
e) Repeat this simulation analysis for 20 and then for 50 babies/mothers. (To save some time, reduce the number of repetitions to 10,000.) Also give an estimate of the uncertainty with these approximations. How does the probability appear to change as number of babies/mothers increases?

Example 1-5: Unfinished Game
The mathematical study of probability originated with a series of communication between famous mathematicians Pascal and Fermat in the 1600s. They were figuring out how to solve a problem similar to this: Suppose that Heather and Tom play a game that involves a series of coin flips. They agree that if 5 heads occur before 5 tails do, then Heather wins the game. But if 5 tails occur before 5 heads do, then Tom wins the game. They each agree to pay $5 to play the game, so the winner will make a profit of $5. The first five coin tosses result in: H₁, T₂, T₃, H₄, H₅. Unfortunately, the game is interrupted at that point and can never be finished.

a) Suggest a reasonable strategy for dividing up the $10 between Heather and Tom.

b) Make a guess for the probability that Heather would win the game if it were to continue.

c) List the sample space of all possible outcomes, using similar notation to the coin toss results above. [Hint: There are 10 possible outcomes.]

d) Consider the events H = {Heather wins the game} and T = {Tom wins the game}. Identify the outcomes that comprise each of these events.
e) Is it reasonable to consider these outcomes to be equally likely? Explain.

We will wait until we have learned some more probability concepts (specifically the concept of independence) before we calculate the exact probabilities here. But we could use simulation to determine approximate probabilities.

f) Describe (in detail) how you could use a coin to conduct a simulation to approximate the probability that Heather wins the game.

g) Indicate one fundamental difference between this simulation and the previous ones about random babies and ice cream prices.

Example 1-6: Solitaire
Suppose that every night I play Solitaire on my computer until I win for the first time.

a) Describe the sample space for this random process.

b) How many outcomes are in this sample space?

c) Are these outcomes equally likely? Explain why or why not.

d) What branch of mathematics might be useful for determining probabilities involving random processes with an infinite number of possible outcomes?