Excerpts from *The Decision Making Book*
Chapter 11: Simulation or "Monte Carlo" Technique

11.1 Introduction

We have been studying how one analyzes random variables. If the characteristic we are studying is a random variable, we want **three pieces of information**:

1) The **average** of the R.V., a typical value, a standard to compare against.
2) The **standard deviation** of the R.V., a number that quantifies the variability of the R.V.
3) The **probability distribution** of the R.V., a picture of the shape of the variability.

This information gives us an insight into how the random variable behaves. We have seen that if we know the probability distribution of the R.V., we can easily determine the average and standard deviation of the R.V.

$$\textbf{Average} = E(X) = \sum X\, P(X)$$

$$\textbf{Standard Deviation} = \sqrt{E(X^2) - (E(X))^2}$$

$$\text{where } E(X^2) = \sum X^2 P(X).$$

The key is to determine the probability distribution of the R.V.
This is done in three ways:

1. **Observed Probability**, which is based on observing what happened. The observed percentage of time a given event happened.

$$P(A) = \frac{\text{Observed Number of Times "A" Happened}}{\text{Total Number of Observations}} = \frac{f(A)}{n}$$

The **Empiricist** argued that all **Knowledge** comes from our **Experiences**. The Observed Probability reflects the Empiricists' view of how we acquire Knowledge that helps us explain the behavior of random variables.

2. **Reasoned Out Probability**, which is based on reasoning out what percent of time it will happen.

$$P(A) = \frac{\text{Reasoned Number of Times "A" Will Happen}}{\text{Reasoned Total Possible}} = \frac{n(A)}{n(\text{Total})}$$

The **Rationalist** argued that all **Knowledge** comes from our ability to **Reason**. The Reasoned Out Probability reflects the Rationalists' view of how we acquire Knowledge that helps us explain the behavior of a random variable.

3. **Opinion Probability**, which is based on one's opinion of how often things would happen.

$$P(A) = \frac{\text{Opinion Number of Times "A" Will Happen}}{\text{Pretend Total Possible}} = \frac{n(A)}{n(\text{Total})}$$

You may have studied two theoretical probability models, normal and binomial. They allow us to determine probabilities without actually running an experiment, taking a sample or forming an opinion. We use the observed probability results to check to see if our reasoned out models were or would be appropriate.

The **Law of Large Numbers** says: "If the reasoning is 'good', the **Observed Probability** will get closer and closer to the **Reasoned Out Probability**, as the number of observations gets larger." If this isn't happening, then we must conclude that the Reasoned Out probability is wrong.
The Reasoned Out Probability is supposed to explain or predict what will actually happen and the Observed Probability is simply a summary of what actually did happen.

We can make a similar statement about Opinion Probability: if the opinion is "correct" we expect the Observed Probabilities to get closer and closer to the Opinion Probabilities, as we gather more experiences. If this doesn't happen, our opinion was wrong.

In some problems that have never happened before, it may be impossible, either because of time or cost, to run an experiment or take a sample to gain "experience" to get the **observed** probabilities. Also, the problem may be so complex that it is impossible to **reason out** the probabilities or the consequences of a bad estimate are so grave that people wouldn't use **opinions.** **Simulation**, or what is sometimes called the **Monte Carlo Technique**, gives us a fourth method of determining the probability of an event.

Apollo 13 Used Simulation

Many people have seen the movie or remember the problem with Apollo 13.

The spacecraft malfunctioned and was not able to land on the moon. The question that mission control had to answer was, "At what point in their orbit of the moon should they fire their engines?" They needed to "pull out" of the moon's gravity and get into the earth's gravitational field. If they fired the engines at an incorrect time, they could have gone out into space.

How should the astronaut react to get back safely to the Earth? Could they use:

Experience?	No, this was a new situation, they had no prior experiences. Also, they could **not** send up several spacecrafts to experiment and observe what to do.
Reasoning?	The problem was so complex that they did not believe they could write mathematical equations that would correctly determine the point to fire the engines.
Opinion?	If you were one of the astronauts, would you want mission control to say we **feel** it, fire the engine now?

NO! NO! NO!

They used a technique called **simulation**. They anticipated a need to do may "pretend" rocket firings. They had a program that simulated the path Apollo 13 would take. In doing this, they could experience life without having to live it—this is the basic creed of simulation.

Radiation and Thickness of Lead Shield

When people began working with radiation, a basic question arose:

How thick does the lead shield have to be to protect humans from the radiation?

Could they use:

Experience? No, they shouldn't use humans for this type of experimentation. Unfortunately, there are situations where people were exposed to radiation when their consent had not been given.

Reasoning? To explain how radiation behaves when it hits a lead shield is so complex, mathematical equations have not been written to explain what will happen.

Opinion? Would you want to be in a room with radiation and hear the person say: "my guts" tell me the shield is thick enough?

They used a technique called **simulation.**

Simulation combines many of the aspects of reasoned out probability with observed probability, and, in some cases, opinions.

The unique aspect of simulated probability is that rather than actually running a given experiment we create an artificial experiment that in an **abstract probabilistic way** is equivalent to our original experiment, problem or decision. The discussion that follows will help clarify this point.

Reflection Airplane pilots in training use what is called a **flight simulator**. What do you think the simulator does?

11.2 Random Number Table

Look at the table on the following page. You could think that the table was created by someone (with lots of time on their hands) who put the numbers 0,1, 2, 3, 4, 5, 6, 7, 8, 9 on ten chips and put the chips in a jar. She then selected one chip at random and recorded the result. In this table, a 4 was selected first, the 4 chip was then returned to the jar, the chips were mixed and a second chip was selected. In this case it turned out to be a 1. The woman enjoyed this so much that she spent the evening randomly selecting and recording the digits. The next page shows her results. If you enjoy reading the next page, we encourage you to pick up a book published by the Rand Corporation, a think tank of the U.S. Army. You will find over 200 pages of random numbers—exciting reading for those lonely Minnesota winter nights. Or you can generate a table such as this using Excel.

What is the probability of selecting the "3" chip from the bowl?

$$P(3) = \frac{n(3)}{n(\text{total})} = \frac{1}{10}$$

What is the probability of placing a pencil at random on the random number table and having it hit or be closest to the digit 3?

$$P(3) = \frac{1}{10}$$

In an **abstract probabilistic way** the selecting of a digit from a jar or a random number table are equivalent.

Simulation of Coin Flipping

Reasoned Out Probability:

$$P(\text{Head}) = \frac{n\ (\text{Heads})}{n(\text{Total Outcomes of a Coin})} = \frac{1}{2}$$

Instead of flipping a coin, we select a one digit number at random from the random number table and use the following rule:

Rule: If digit is **even**, a **head** has occurred.
If digit is **odd**, a **tail** has.

$$P(\text{Even number}) = \frac{n(\text{Even Digits})}{n(\text{Total})} = \frac{5}{10} = \frac{1}{2}$$

We have **simulated** coin flipping, because in an **abstract probabilistic way**, the flipping of a coin and the selecting of an even digit can be considered equivalent events. The probabilities are equal. Instead of actually flipping a coin, we **simulate** the experiment by selecting a digit at random from the table.

Table of Random Digits

4 1 2 5 5 9 2 2 9 2 4 2 7 9 2 4 7 0 4 4 0 8 8 8 7 5 3 4 6 2 2 7 0 6 1 9 1 1 2 4
3 0 0 8 2 1 0 6 7 3 9 3 6 0 0 9 7 1 6 1 3 5 9 2 9 0 3 9 3 0 7 8 1 3 8 2 9 0 0 6
9 5 1 1 5 7 9 1 1 8 6 6 2 3 9 8 2 0 7 1 8 4 4 8 7 5 4 2 3 7 3 0 1 5 9 9 9 5 6 9
9 0 9 6 3 5 2 6 5 8 4 4 2 6 5 0 1 8 6 7 2 1 6 3 5 3 1 3 5 2 5 0 2 2 0 5 4 2 1 8
8 1 6 1 0 8 1 4 8 2 6 2 7 4 7 1 1 4 5 2 6 4 7 8 5 7 6 5 7 1 1 2 5 9 2 3 7 4 2 1
5 1 1 3 0 0 1 1 8 2 2 8 8 4 1 6 3 9 2 5 1 6 9 8 7 4 5 4 5 0 0 3 0 2 4 2 4 8 3 0
0 3 1 9 1 3 9 2 6 9 7 2 1 4 6 4 7 6 2 2 3 2 3 0 5 0 7 0 8 8 4 7 4 4 3 8 6 3 1 3
7 9 8 2 7 0 0 8 0 1 3 6 3 6 9 9 5 2 3 3 3 7 7 2 6 9 3 5 6 7 5 1 8 1 4 5 4 8 7 6
3 3 7 0 1 5 4 5 7 2 0 5 2 5 7 0 7 8 6 2 4 4 8 6 1 9 5 7 4 7 3 2 7 0 5 1 3 1 6 3
9 4 2 3 1 1 9 1 9 8 1 1 3 5 8 3 1 8 5 7 9 7 7 1 4 1 2 9 1 1 3 6 5 6 5 2 0 7 5 7
9 2 0 7 3 2 6 3 9 5 8 7 2 7 5 9 4 6 6 9 2 8 6 2 6 6 1 5 4 7 7 1 3 2 2 5 2 3 1 8
1 4 4 2 1 1 2 8 1 6 8 3 6 6 3 3 5 3 0 2 5 0 0 7 5 1 3 1 9 5 1 1 7 7 0 5 6 1 3 9
0 4 4 0 7 8 2 3 4 0 3 0 4 3 9 0 5 1 7 6 8 6 6 4 3 9 7 7 8 9 9 9 3 2 3 9 4 5 4 5
0 2 7 4 5 4 1 7 3 3 3 8 4 5 7 1 2 9 6 7 0 0 3 9 3 0 9 8 6 5 0 6 0 3 4 9 2 2 4 0
0 0 6 2 0 7 7 9 4 1 7 8 6 6 9 8 8 5 8 5 0 2 0 8 5 4 7 8 4 6 0 0 1 2 4 8 8 7 1 5
5 4 3 7 9 7 7 5 4 9 6 2 9 8 8 9 8 0 7 4 4 1 3 2 6 0 9 2 3 2 6 4 6 3 5 3 1 9 4 5
6 0 3 2 8 2 2 4 2 3 9 0 8 5 6 2 2 2 7 5 0 7 7 8 0 5 2 5 7 9 4 7 6 1 6 9 0 1 0 9
0 9 8 9 4 4 3 0 0 1 0 5 6 7 1 2 5 9 7 6 6 4 4 2 6 0 7 7 4 7 2 3 8 0 5 3 6 2 0 4
3 7 5 2 5 9 9 6 9 0 6 3 0 9 4 6 3 5 1 0 3 5 0 7 0 8 2 4 9 0 2 2 8 9 2 7 5 3 0 5
2 4 3 6 6 7 3 7 0 2 3 9 2 5 7 6 4 3 9 0 8 3 9 7 5 6 9 5 2 1 9 4 4 3 0 7 0 0 8 2
6 7 4 4 5 5 3 3 9 0 3 7 9 4 1 8 7 8 5 3 7 9 3 3 1 7 6 9 2 5 4 4 9 5 3 6 6 7 9 0
2 9 0 2 5 4 1 8 8 5 8 0 0 2 5 7 3 4 0 5 7 7 7 2 2 0 0 4 8 7 5 9 3 3 3 6 8 7 9 4
6 5 5 4 3 6 1 1 7 4 0 4 2 1 5 6 5 3 6 9 0 0 8 4 1 7 1 8 5 4 6 2 9 2 0 7 0 5 0 4
1 5 9 4 6 0 4 4 3 4 8 4 1 4 1 3 3 6 2 6 3 5 1 9 4 7 2 7 9 6 3 8 6 4 7 1 0 9 4 5
0 1 9 3 8 1 4 2 9 8 4 4 9 4 0 7 6 9 0 6 7 8 0 0 7 2 6 7 5 8 9 0 1 2 6 8 1 2 4 7
3 6 3 4 5 3 2 6 9 7 6 8 9 3 2 4 9 1 1 5 2 5 6 5 5 1 2 6 1 9 7 6 2 3 3 7 6 1 2 1
8 7 7 2 8 0 0 2 4 4 6 2 7 5 3 9 4 6 4 1 8 2 9 3 0 1 1 0 6 0 8 4 4 9 0 6 6 3 2 4
9 8 9 2 7 6 9 4 2 8 0 5 5 2 6 4 9 9 5 0 5 8 3 2 1 4 7 4 6 4 9 8 4 5 7 4 6 6 9 1
5 8 4 0 5 7 5 8 3 8 9 5 7 4 0 5 9 7 8 8 6 1 1 7 7 6 6 4 4 7 7 0 4 0 5 8 5 9 6 7
4 8 9 9 3 2 9 0 2 2 7 6 5 1 2 3 9 7 1 3 8 9 3 8 9 7 3 4 7 1 4 0 1 2 1 9 7 9 3 0
8 1 1 1 4 1 8 7 7 7 8 2 5 1 7 0 5 6 9 5 0 0 5 2 7 7 8 7 4 8 1 2 8 0 7 5 4 5 6 6
2 7 1 5 0 3 9 9 3 2 2 1 1 3 3 2 5 4 1 8 5 2 4 0 7 7 7 7 4 5 3 2 1 4 3 5 0 3 7 1
4 5 9 4 2 8 5 9 2 2 3 0 5 0 2 4 9 1 0 4 3 9 7 4 5 4 7 0 9 9 7 4 9 7 5 4 3 9 8 1
3 5 3 3 1 7 3 6 2 9 6 8 8 2 0 0 1 5 7 1 7 4 7 6 1 7 1 3 5 5 2 7 4 8 3 1 1 8 7 4
9 8 1 3 2 2 4 4 4 6 2 9 7 9 4 0 1 7 5 4 8 5 0 7 9 7 2 1 8 3 5 7 0 0 5 2 3 9 5 5
8 9 1 3 1 5 1 8 4 7 0 2 5 7 7 8 4 2 9 5 7 0 2 6 3 7 5 9 8 8 3 5 2 9 9 8 2 0 9 5
2 4 0 6 0 3 5 3 6 6 2 6 3 5 8 1 3 5 4 1 6 1 1 1 9 2 9 1 3 3 0 6 9 9 1 5 5 0 0 0
5 2 2 6 5 3 9 2 2 9 3 2 2 0 7 3 6 8 1 7 4 3 8 2 0 4 5 0 3 6 7 2 0 8 5 6 6 0 4 5
4 2 3 3 5 0 0 9 4 1 7 9 2 5 0 2 9 5 8 0 8 4 2 1 9 2 7 6 2 0 2 9 4 4 1 8 8 6 6 3
1 1 1 6 2 1 9 3 9 9 3 5 4 4 3 4 6 1 1 6 9 4 8 8 9 2 4 2 3 4 0 2 2 7 7 8 2 6 5 7

The table was generated using Excel.

Problem. Flip a coin 12 times and use these results to determine the probability of flipping a head.

Numbers Selected At Random From Table

3 = tails	0 = heads
6 = heads	3 = tails
9 = tails	2 = heads
6 = heads	0 = heads
5 = tails	3 = tails
0 = heads	4 = heads

Instead of actually flipping a coin twelve times, we simulated the experiment by using the random number table. The result for this simulation is seven heads and five tails. The simulated probability of a head would be 7/12, which is close to 1/2. We'd expect that the **Simulated Probability** would get closer to the **Reasoned Out Probability** as the number of simulated trials increases, we have faith in the **Law of Large Numbers.**

Simulation of Poker

Instead of dealing five cards from a deck of cards, we could put 52 different numbers into a jar and pick five of them out. Then we use the following table to determine what cards were "dealt". In simulating any 52 different numbers could have been used. We could have used the numbers 121 to 172. Instead of putting the numbers in a jar, we could select a 2 digit number from the random number table. We would select five two digit numbers at random that are between and including 00 to 51. If we select a number twice, we'd select another number as we can't draw the same card twice. If we get a two digit number greater than 51, we'd select another two digit number.

RN	Clubs	RN	Diamonds	RN	Hearts	RN	Spades
00	Ace	13	Ace	26	Ace	39	Ace
01	King	14	King	27	King	40	King
02	Queen	15	Queen	28	Queen	41	Queen
03	Jack	16	Jack	29	Jack	42	Jack
04	10	17	10	30	10	43	10
05	9	18	9	31	9	44	9
06	8	19	8	32	8	45	8
07	7	20	7	33	7	46	7
08	6	21	6	34	6	47	6
09	5	22	5	35	5	48	5
10	4	23	4	36	4	49	4
11	3	24	3	37	3	50	3
12	2	25	2	38	2	51	2

Simulation involves describing actual events or outcomes as numbers and the probability of selecting a number is equal to the probability of the event the numbers are depicting.

$$\textbf{P(Selecting an Ace of Hearts)} = \frac{\textbf{n(Ace of Hearts)}}{\textbf{n(Total Cards)}} = \frac{1}{52}$$

$$\textbf{P(Selecting the number 26 from 00 to 51)} = \frac{\textbf{n(Pick the number 26)}}{\textbf{n(Total Numbers)}} = \frac{1}{52}$$

n(26) = 1, because the number 26 occurs once in the list of two digit numbers that go from 00 to 51.

Again, in an **abstract probabilistic way** these two events, selecting an Ace of Hearts and selecting the number 26, are equivalent.

Reflection Do you understand what a table of random numbers is?

11.3 Computer Simulation

The primary reason simulation has become more important is that it can be done easily on the computer.

Rather than having to pick numbers at random from a table or bowl, we can program the computer to, in essence, select numbers at random from a jar.

Most computers have as part of their software package a **random number generator**.

This is a program of mathematical equations that will create (generate) numbers that are random in nature.

You may have seen or heard about the fight of the century, where Mohammed Ali fought Rocky Marciano.

This fight never happened, it was a simulation. Other simulations have been done trying to guess which was the greatest baseball team of all time.

You must remember that these results prove nothing and are very much subject to the simulator's whims and biases.

However, many practical problems are done using simulation and you shouldn't be swayed by a few applications that deal with sensationalism.

We will look at two computer software packages, **@Risk** and **Crystal Ball.** They both allow you to have your spreadsheet do simulation.

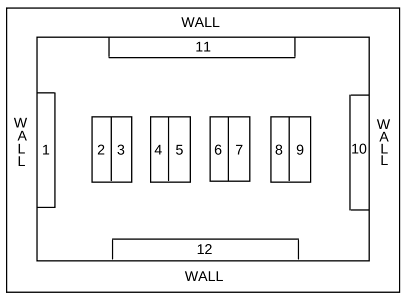
We will look at a few problems without using the computer prior to looking at **the software**, as it is important for you to understand what the computer does for you.

Random Samples

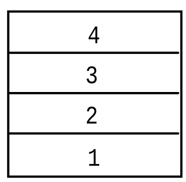
In statistics we learn about the desire to take samples that are random in nature—that is, every object in the population would have an equal chance of being selected. This can be very difficult to achieve in real life problems. We look at a method of using random numbers to help ensure a random sample.

Problem: Price Comparisons
We wish to compare the prices of two stores and determine which is the cheapest. How do we go about taking a random sample of the items in the store(s)? If we just go in and pick items out, **our biases** may lead us to select certain items and overlook other items. If we do this we may have made a **sampling error**, the sample does **not represent** the **population**, all the items in the store.

The following is a possible way of getting a sample that would be nearly random.
We make a layout of one of the stores and number the aisles.



If we assume that there are 4 shelves on each aisle, by picking two numbers at random we could select the aisle we want and the shelf.



"7-3" would indicate we want an item from aisle 7 on shelf 3. To specify a particular item on the shelf we might pick a number at random between 1 and 50 and then count from the front to back and returning until we reach our desired number (the 50 is completely arbitrary).

"10-2-28" would be the item found in the 10th aisle on shelf 2 and the 28th item using the above numbering scheme. We would pick a number at random from a bowl that contains numbers 1 through 12, to give us the aisle. A second number from a bowl containing numbers from 1 through 4 will give us the shelf, and a third number from a bowl containing numbers from 1 through 50 would give us the item.

Fortunately, we don't have to create bowls, we can use computer programs that will generate random numbers. The following page indicates output if we had wanted a random sample of size 20 (n = 20). Note that the numbers are arranged in ascending order. They were not generated this way. The first sequence of numbers generated was **8 1 11** , but by having them in this order it is easier to gather the data.

Using the table I'd go to aisle 1, shelf 1 and find the 28th item, etc.
The program is menu driven and we show the questions you must answer to get the desired output. What is underlined is what we had to type in. The program is called RASEQ, which is short for Random Sequence.

Do you need instructions? **Yes**

RASEQ generates up to 250 sets of 1 to 4 random numbers.

Suppose, for example, that you wanted to select 100 items from a 500-page computer listing with 60 items on each page. You would want 100 sets of 2 numbers, the first number in the range 1 to 500 and the second in the range 1 to 60. The program will ask for the number of sets you want, n, and the quantity of items to be included in each set, q. The L and H values are higher and lower limits.

Enter the values of n and q? **20,3** ---------------- sample size

For the first number -- L, H = ? **1,12** ---------------- aisles go from 1 to 12

For the second number -- L, H = ? **1,4** ---------------- shelves

For the third number -- L, H = ? **1,50** ---------------- Items

Order Of Selection		**Set Of Random Numbers**		
3		1	1	28
15		2	1	25
18		2	3	32
4		4	4	6
10		4	4	20
19		5	1	34
7		5	4	18
8		6	2	13
6		7	2	1
13		7	2	40
9		7	3	15
1	--- First sequence	8	1	11
14	of numbers that	8	3	27
20	the computer	9	2	25
12	generated	9	2	36
2		9	4	17
16		10	1	42
5		10	3	33
17		11	1	38
11		12	2	13

Reflection Can you think of a problem at work that you could simulate?

11.4 Simulation: Cost of a Guarantee

A small company offers to fly people into Canada for fishing trips. They charge $400 for the pilot, plane and fuel. The customer must canoe out so that there is no charge for pick up (makes problem much easier).

Assume that when customers call for appointments, they are told there can be no guarantee on the departure day but they will get $50 for each day that they must wait.

The company wishes to determine the cost associated with this guarantee. They must keep the guarantee as their competitor has this kind of guarantee. When they did not offer the guarantee, their number of customers dropped dramatically.

Can you identify the two random variables of this problem?

1) The number of fishing parties **arriving** each day that want service is a random variable.
2) The number of flights that they are able to make each day is a random variable. This is the same as the number of fishing parties they are able to **service** each day, it is also a random variable.

Do you see the problem? On some days 3 parties may want a flight and because of weather conditions, equipment problems, etc., the company is only able to "service" 2 of the customers. Hence, there is one party waiting.

To analyze how much waiting will take place, we must look at the two random variables:

Arrivals and Service.

From their database they have the following information on arrivals and service.

They Believe the Past Will Mirror the Future.

Arrivals = Number of Parties seeking a plane trip each day
Service = Number of Parties who are flown out each day

Arrivals	P (Arrival)	Potential Services	P (Service)
0	.15	0	.05
1	.20	1	.10
2	.25	2	.20
3	.30	3	.20
4	.10	4	.25
		5	.20

These would be **Observed Probability Distributions**. To reason out how these two random variables interrelate to determine the average waiting time, requires extensive mathematics. Rather than attempting to **reason out** the solution, we **simulate** the arrival and service of customers using the random number table.

The following table indicates the relationship that exists between the random numbers selected and what they represent in terms of service and arrival values.

Arrivals	Random Number	Service	Random Number
0	00-14	0	00-04
1	15-34	1	05-14
2	35-59	2	15-34
3	60-89	3	35-54
4	90-99	4	55-79
		5	80-99

Note how in an **abstract probabilistic way** the selecting of the numbers 00 to 14 from a random number table

$$\textbf{P(Pick 00 to 14)} = \frac{\textbf{n(2 digit numbers from 00 to 14)}}{\textbf{n(Total number of 2-digit numbers)}} = \frac{15}{100} = .15$$

is **equivalent** to what has been observed for parties arriving:

$$\textbf{P(No customers arrive)} = \frac{\textbf{f(No parties seeking service)}}{n} = .15$$

Day (1)	RN for Arrival (2)	No. of Arrival (3)	No. of Parties Left from Yesterday (4)	Total No Parties Waiting (5)	RN for Service (6)	No. of Parties served Today (7)	No. of Parties Left Unserved (8)	Cost (9)
1	73	3	0	3	21	2	1	50
2								

ETC.

We explain Line 1. First we randomly selected a two digit number from the Random Number Table; the number selected was 73. We looked in the table above on arrivals and saw that 73 falls in the range of random numbers that indicates 3 arrivals. The 0 in the fourth column comes from assuming there is no one waiting. Column 5 is 3 + 0 = 3. There were 3 new arrivals plus no one waiting from yesterday.

We then randomly picked another two digit number, and it was 21. We checked in the table above on number serviced and found this corresponds to 2 parties being serviced. This leaves one (3 - 2=1) party that must wait. This will cost the company $50. Using a computer, we could have the computer do this one thousand times and see if patterns appeared in "long lines" of people waiting, and what it would cost the company.

Reflection How did this example help you think of another problem at work you could simulate?