I. Chapter 1

Introduction to Probability Theory

I.1. Case Study

(Motivation: Application of the topics that will be covered in this chapter to the real world problems)

Radelet (1981) studied effects of racial characteristics on whether individuals convicted of homicide receive the death penalty. The events that are considered on this study are the selection of a case with "death penalty verdict", "not death penalty verdict", "white defendant", "black defendant", "white victim", and "black victim". The 326 subjects were defendants in homicide indictment in 20 Florida counties during 1976-1977. The following table gives the number of subjects for each of the defendant's race, victim's race and death penalty combinations.

		Death Penalty		
Defendant's Race	Victim's Race	Yes	No	Total
White	White	19	132	151
	Black	0	9	9
Black	White	11	52	63
	Black	6	97	103
Total		36	290	362

Source: Agresti, A. Categorical Data Analysis, John Wiley & Sons, 1990, pg. 135-138

The main question that one would like to answer is "Is there an evidence of racial discrimination given the evidence on this table?". Also, one would be interested with the following questions; (i) is there a relation between defendant's race and victim's race?, (ii) is there a relation between victim's race and death penalty? (iii) If we control for the victim's race, that is if we look at the cases for black victims and white victims separately, what is the relation between defendant's race and death penalty verdict?

I.2. Historical Remarks
(Motivation)

In the long run, we are all dead.
-John Maynard Keynes (England, 1883-1946)

People talk loosely about chance all the time, without doing any harm. What are the chances of getting a job? of meeting someone? of rain tomorrow? of passing Math. 3610? But for scientific purposes, it is necessary to give the word *chance* a definite, clear interpretation. This turns out to be hard, and mathematicians have struggled with the job for centuries. They have developed several careful and rigorous theories of chance; but the theories cover just a small range of cases where people ordinarily speak of chance.

Exercise 1. How would you explain the fact that when you toss a coin the chance of getting head is one-half? How many possibilities we have if you toss a coin? Can you repeat the experiment, that is tossing a coin, under the same conditions many times?

Exercise 2. What is the meaning of "chance of passing Math. 3610 is 0.80"? How many possibilities we have for this experiment? Can you repeat the experiment, that is taking Math. 3610, under the same conditions many times? What could be the conditions that would change if you repeat the experiment?

Frequentist definition of chances works best for the processes that can be repeated over and over again, independently under the same conditions. Games of chance fall into this category, and in fact much of the frequency theory was developed to solve gambling problems. One of the great masters was Abraham de Moivre, a French Protestant who fled to England to avoid religious persecution. In his book *The Doctrine of Chances* he included the following letter,

> To the Right Honorable the
> Lord CARPENTER
>
> My Lord
>
> There are many people in the World who possessed with an Opinion, that the Doctrine of Chances has a Tendency to promote Play; but they soon will be undeceived... Your Lordship does easily perceive, that this Doctrine is far from encouraging Play, that is rather a Guard against it, by setting in a clear light, the Advantages and Disadvantages of those Games wherein Chance is concerned...
>
> Another use to be made of this Doctrine of Chances is that it may serve in conjunction with the other parts of the Mathematicks, as a fit introduction to the Art of Reasoning: it being known by experience that nothing can contribute more to attaining that Art, than the consideration of a long Train of Consequences, rightly deduced from undoubted Principles, of which this Book affords many examples.

Historically, probability had its origin in the gambling room. The Chevalier de Mere, a professional French gambler, had asked his friend Blaise Pascal (1623-1662) to solve the following problem: In what proportion should two players of equal skill divide the stakes remaining on the gambling if they are forced to stop before finishing the game? Pascal wrote to and began an active correspondence with Pierre Fermat (1601-1665) concerning the problem (collaboration). Although Pascal and Fermat agreed on the answer, they gave different proofs (nonuniqeness of the solutions). It was in this series of correspondences during the year 1652 that they developed the modern theory of probability.

A century earlier the Italian mathematician and gambler (interesting combination!) Girolomo Cardan (1501-1576) wrote *The Book on Games of Chances*. This is really a complete textbook for gamblers since it contains many tips on how to cheat successfully. Cardan was also an astrologer. According to the legend, he predicted his own death astrologically and to guarantee its accuracy he committed a suicide on that day. (Of course, that is the most convincing way to be right!) Like Cardan, De Moivre predicted the day of his death. A rather interesting story is told of De Moivre's death. De Moivre was ill and each day he noticed that he was sleeping 15 minutes longer than he did on the preceding day. Using progressions, he computed that he would die in his sleep on the day after he slept 23 hours and 45 minutes. On the day following a sleep of 24 hours, De Moivre died.

The French mathematician Pierre Simon de Laplace (1749-1827) also contributed much to the historical development of probability. In 1812 he published *Theorie Analtique des probabilities,* in which he referred to probability as a science that began with games but that had wide-ranging applications. In particular, he applied probability theory not to gambling situations but as an aid in astronomy.

Over the course of many years probability theory has left the gambling rooms and has grown to be an important and ever-expanding.

I.3. Review

Set Theory

In this section we will look at the basic notions and notations needed for dealing with sets (collection of objects of various kinds, such as numbers or paired of numbers, called points). Although you may have dealt with these concepts since grade school, it is important that this chapter be utilized to assure the sound base for the applications of the probability that follow in Chapter 1 of your textbook.

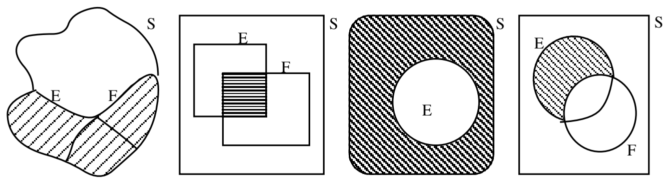
"Set" is an undefined notion (like "point" or "line" in high-school geometry); it is assumed that at least one exists, and attention is restricted to a universe or universal set. In probability theory universal set corresponds to the sample space S, sets correspond to the events. All operations on sets are with respect to S in order to avoid paradoxical situations.

Exercise 4. The Barber of Seville shaves all men who don't shave themselves. Who shaves the Barber of Seville.

Basic Set Operations

There are four basic set operations which we will illustrate their meanings with diagrams called Venn diagrams.

Definition 1. Union (or **join**, or **sum**) of sets E and F:
$E \cup F \equiv E + F \equiv \{w: w \in E \text{ or } w \in F\}$ (which you read the last notation as E union F consists of points w which are elements of E or F).



$E \cup F$ $E \cap F$ E^c $E - F \equiv E \cap F^c$

Definition 2. Intersection (or **meet**, or **product**) of sets E and F:

$$E \cap F \equiv EF \equiv \{w: w \in E \text{ and } w \in F\}.$$

Definition 3. Complement of set E:

$$E^c \equiv \bar{E} \equiv E' \equiv \{w: w \in S \text{ and } w \notin E\}.$$

Definition 4. Difference of sets E and F:

$$E - F \equiv E \cap F^c \equiv \{w: w \in E \text{ and } F \notin F\}.$$

Additional Definitions on Sets

Definition 5. E is a **subset** of F, written $E \subset F$, if $w \in E \Rightarrow w \in F$.
Definition 6. The **empty set** (or **null set**), usually written $\emptyset$, is the set containing no points at all.
Definition 7. Sets E and F are said to be **mutually exclusive** (or **disjoint**) if
$E \cap F = \emptyset$.

Simple Laws and Some Extensions on Sets

In the following theorem we are going to state some simple laws dealing with the operations on sets.

Theorem on Sets: Let S be the sample space and E, F, and G be the events on S. Then
1. $E \cap E = E$, and $E \cup E = E$ (**Idempotency Laws**).
2. $E \cap F = F \cap E$, and $E \cup F = F \cup E$ (**Commutative Laws**).
3. $E \cap (F \cap G) = (E \cap F) \cap G$, and $E \cup (F \cup G) = (E \cup F) \cup G$ (**Associative Laws**).
4. $E \cap (F \cup G) = (E \cap F) \cup (E \cap G)$, and $E \cup (F \cap G) = (E \cup F) \cap (E \cup G)$ (**Distributive Laws**).
5. $E \cap S = E$, $E \cup S = S$,
$E \cap \emptyset = \emptyset$, $E \cup \emptyset = E$,
$E \cap E^c = \emptyset$, $E \cup E^c = S$,
6. $(E \cap F)^c = E^c \cup F^c$, and $(E \cup F)^c = E^c \cap F^c$ (**De Morgan's Laws**) .
7. $(E^c)^c = E$
8. F is a subset of E if and only if $E \cap F = F$.
9. F is a subset of E if and only if $E \cup F = E$.
Although the terminology given in parentheses is not now in widespread use, it is important to expose to it in order to be able to read other (and older) books and writings.

Exercise 5. Translate the following informal statements and definitions into the formal form by using the set notations.

Occurrence of two thing at the same time	
Occurrence of one thing prevents the occurrence of the other	
Occurrence of one thing implies the occurrence of the other	
Occurrence of either one of the two thing	
Occurrence of an opposite of one thing	
Occurrence of three things at the same time	
Occurrence of only one of the two things	
Occurrence of one thing given the occurrence of the other	
Occurrence of any one of the possibilities	
Occurrence of an neither one of the two things	
Occurrence of the either one of the three distinct things	

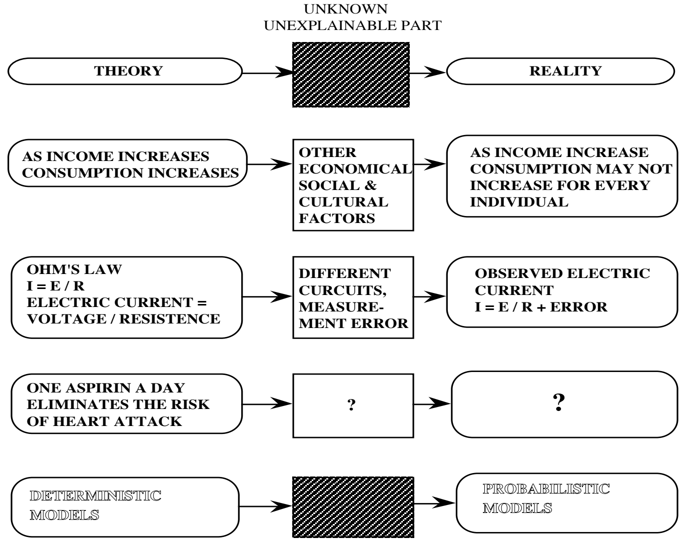
I.4. Overview

Where We're Going

Today, probability theory is a well established branch of mathematics that finds its applications in every area of scholarly activity from music to physics, and in daily experience from weather prediction to predicting the risk of a new medical treatment.

Most important of all it will give us the necessary tools to find a measure of reliability for the generalizations that we are going to make from sample to the whole population. For example when you say "Vikings will win the super bowl in 1992-93", is this statement reliable 100%? What is the reliability measure of this statement?

To understand the importance of the probability theory better, we should grasp the difference between *theory* and *reality*. But note that reality does not always turn out to be same as theory, that is what we expect. For example, in the theory if you get your graduate degree, you should be able to get a well-paid job. But a person with a graduate education may end up with a low-paid job.



Theories are ideas proposed to explain phenomena in the real world and, as such, are approximations or models for reality. These models are expressed either in a verbal form or as mathematical relationships. Whereas the theory of social change might be expressed verbally in sociology, the theory of heat transfer is presented in a precise and deterministic mathematical manner in physics. Neither gives an accurate and unerring explanation of the real life. Slight variations from the mathematically expected can be observed in heat transfer phenomena and other areas of physics. The deviations can not be blamed solely on the measuring instruments, the explanation that one often hears, but are due in part to a lack of agreement between theory and reality. Anyone who believes that the physical scientist completely understands the wonders of this world need only look at history to find a contradiction. Theories assumed to be the "final" explanation for nature have been superseded in rapid succession during the past century.

In this course, we shall develop certain models of reality (probabilistic models); we shall attempt to explain the motivation behind such a development and uses of the resulting model. We will never claim that some statement is always correct, but we will attach reliability, a type of chance statement on the validity of our statement.

Reading Probabilistic Statements

Exercise 6. Translate the following informal statements and definitions into the formal form by using the notations used in your text. The informal statements are taken from the *Doctrine of Chances*.

The chance of something gives the percentage of time it is expected to happen, when the basic process is done over and over again, independently under the same conditions.	P(E)=
The chances are between 0% and 100%	
The chance of something equals 100% minus the chance of the opposite thing.	

If the occurrence of one thing prevents the occurrence of the other, then the chance of observing these two things together is zero.	
If the chance of occurrence of one thing is not affected by the occurrence of the other thing they are called independent.	
Given the occurrence of one thing the chance of occurrence of another thing is proportional to chance of observing these two together, and inversely proportional to the chance of observing the given thing	
The chance of observing two things at the same time is equal to the chance of observing one of them times the chance of observing other given the one we know its occurrence chance.	
Two things are independent if the chances for the second one given the first are the same, no matter how the first one turns out.	
The chance of occurrence of one thing is equal to the chance of occurrence of a second thing and that thing together, plus the chance of nonoccurrence of the second thing and that thing together.	
If two things can not occur together than the chance of observing either one of them is the sum of chance of observing each.	

I.5. Technical Details

The following are the basic results that you have learned in this chapter:

$P(E^c) = 1 - P(E)$
$P(E \cup F) = P(E) + P(F) - P(E \cap F) = P(E) + P(F) - P(EF)$
$P(E \mid F) = \frac{P(EF)}{P(F)}$

If E and F are mutually exclusive then, $P(EF) = 0$
If E and F are independent then, $P(EF) = P(E)P(F)$
If E and F are independent then, $P(E \mid F) = P(E)$ and $P(F \mid E) = P(F)$

$P(E) = P(EF) + P(EF^c)$

Exercise 7. Fill out the following table with the formulas that you would use to find the required probability given the information in the first column.

GIVEN	FIND	GENERAL	MUTUALLY EXCLUSIVE	INDEPENDENT
P(E), P(F), $P(E \cup F)$	P(EF)	$P(EF) = P(E) + P(F) - P(E \cup F)$	$P(EF) = 0$	$P(EF) = P(E)P(F)$
$P(E \mid F)$, P(F)	P(EF)			
$P(F \mid E)$, P(E)	P(EF)			
$P(E \mid F)$, P(F), P(E)	P(EF)			
$P(E \mid F)$, $P(E \mid F^c)$, P(F)	P(E)			
$P(E \mid F)$, $P(E \mid F^c)$, P(F)	$P(F \mid E)$			
P(E), P(F), P(G), P(EF), P(EG), P(FG), P(EFG)	$P(E \cup F \cup G)$			
P(EFG), P(FG)	$P(E \mid FG)$			
P(EFG), P(E)	$P(FG \mid E)$			

Independent, Mutually Independent and Pairwise Independent Events

Definition: The events $E_1, E_2, \ldots, E_n$ are said to be <u>mutually independent</u> if

$$P(E_1 \; E_2 \; \ldots \; E_n) = P(E_1)P(E_2) \ldots P(E_n).$$

Definition: The events $E_1, E_2, \ldots, E_n$ are said to be <u>pairwise independent</u> if

$$P(E_i \; E_j) = P(E_i)P(E_j) \text{ for all } i \neq j.$$

Definition: The events $E_1, E_2, \ldots, E_n$ are said to be <u>independent</u> if for every subset $E_{1'}, E_{2'}, \ldots, E_{r'}$, $r \leq n$, of these events

$$P(E_{1'} \; E_{2'} \; \ldots \; E_{r'}) = P(E_{1'})P(E_{2'}) \ldots P(E_{r'}).$$

Exercise 1: Consider the set of events E_1, E_2, E_3, E_4. What properties need to be shown to establish that these events are independent?

For random events we consider the notions of mutually independence and pairwise independence. The questions of interest are: (i) Are there sets of random events which are pairwise independent but not mutually independent?, (ii) Conversely, are there sets of random events which are mutually independent but not pairwise independent? In the following two examples we will try to answer these questions.

Example (pairwise independence does not imply mutual independence): (Berstein, 1928) Suppose that a box contains 4 tickets labeled by 112, 121, 211, and 222. Experiment consists of drawing one ticket at random from the box. Let

E_1 = "1 occurs at the first place"
E_2 = "1 occurs at the second place"
E_3 = "1 occurs at the third place".

It follows that

$P(E_1) = P(E_2) = P(E_3) = \frac{2}{4} = \frac{1}{2}$,

$P(E_1 \; E_2) = P(\{112\}) = \frac{1}{4} = P(E_1)P(E_2)$,

$P(E_1 \; E_3) = P(\{121\}) = \frac{1}{4} = P(E_1)P(E_3)$, and

$P(E_2 \; E_3) = P(\{211\}) = \frac{1}{4} = P(E_2)P(E_3)$.

That is the events E_1, E_2, E_3 are pairwise independent. However the event $E_1 E_2 E_3 = \varnothing$. Therefore

$P(E_1 E_2 E_3) = 0 \neq P(E_1)P(E_2)P(E_3) = \frac{1}{2^3} = \frac{1}{8}$.

Exercise 2: A box contains eight tickets, each labeled with a binary number. Two are labeled 111, two are labeled 100, two 010, and two 001. An experiment consists of drawing one ticket at random from the box. Let E_1 be the event "the first digit is 1", E_2 be the event "the second digit is one", and E_3 the event "the third digit is 1". Show that E_1, E_2, and E_3 are pairwise independent but *not* mutually independent.

Example (mutual independence does not imply pairwise independence):
An experiment consists of tossing two different standard dice, white and black. The sample space S of the outcomes consists of all ordered pairs (ij), i,j = 1, 2, ..., 6, that is

$$S = \{11, 12, 13, ..., 66\}.$$

Each point in S has a probability 1/36. Define the following events:

E_1 = "first die is 1, 2 or 3"
E_2 = "first die is 3, 4 or 5"
E_3 = "sum of the faces is 9".

Therefore,

$E_1 E_2 = \{31, 32, 33, 34, 35, 36\}$,
$E_1 E_3 = \{36\}$,
$E_2 E_3 = \{36, 45, 54\}$,
$E_1 E_2 E_3 = \{36\}$.

It follows that

$P(E_1) = P(E_2) = \frac{1}{2}$, $P(E_3) = \frac{1}{9}$ and

$P(E_1 E_2 E_3) = \frac{1}{36} = \frac{1}{2}\frac{1}{2}\frac{1}{9} = P(E_1)P(E_2)P(E_3)$.

However the events are *not* pairwise independent, because

$P(E_1 E_2) = \frac{1}{6} \neq \frac{1}{2}\frac{1}{2} = P(E_1)P(E_2)$,

$P(E_1 E_3) = \frac{1}{36} \neq \frac{1}{2}\frac{1}{9} = P(E_1)P(E_3)$,

$P(E_2 E_3) = \frac{1}{12} \neq \frac{1}{2}\frac{1}{9} = P(E_2)P(E_3)$.

Exercise 3: In exercise 2, let us change the number on one ticket from 111 to 110, and the number of another ticket from 100 to 101. Show that E_1, E_2, and E_3 are, this time, mutually independent but *not* pairwise independent.

How to Solve Problems Related with the Probability?

Most important element in the solution of the probability and statistics problems is the nonuniqueness of the solution method. The following graphical and tabular techniques are aimed to help to understand the problem better:

I. Straight forward logical set up,
II. Tables representations,
III. Venn diagrams,
IV. Trees.

Straight forward logical set up gives the fastest solution of the problem. But it requires a deep understanding of the definitions, concepts, theorems, and formulas. Last three techniques help to eliminate this difficulty. These techniques help to develop a logical and systematic thinking process on the solution of the problem. Direct, blind-folded use of formulas might endanger logical thinking and generally leads to an incorrect solution of a problem. Whatever technique one prefers, the basic steps are;

(i) Define the experiment,
(ii) Define the sample space (simple outcomes of the experiment)
(iii) Define the events that are given in the problem by using capital letters
(iv) For compound events use the set notation that we discussed before
(v) Write down the event that we would like to find the chance of occurrence on
(vi) Select the formula that will give us that probability given the information on the events you defined in (iii) & (iv).
(vii) Use your intuition and prior knowledge on the result and interpret that number. Does it make any sense? Is it what you expected?, If not, did you carry out your calculations correctly?

Exercise 8: A softball team has three pitchers, A, B, and C, with winning percentages of 0.4, 0.6, and 0.8, respectively. These pitchers pitch with frequency 2,3, and 5 out of every 10 games, respectively.
(a) What is the probability that the team will win the game?
(b) Suppose that the team won the game what is the probability that A pitched the game?

Understanding the problem:
What is the experiment? : Each game in this problem is the experiment
What are the possible outcomes of the experiment? : First of all A may pitch the game or B may pitch the game or C may pitch the game. So let us label each one of these simple events

A = the event that A will pitch the game
B = the event that B will pitch the game
C = the event that C will pitch the game

Secondly, the team may win the game or lose the game, so let us label these outcomes

W = the event that the team will win the game
L = the event that the team will lose the game

What is the relation between the events? That is are they mutually exclusive or independent?

Since the team cannot win and lose the game at the same time W and L are mutually exclusive.

If we assume that only one pitcher pitches the whole game A, B, and C are also mutually exclusive.

The questions that you must have at this point are the following:
Are the events A, W mutually exclusive? If yes, what does it say for the pitcher A?
Are the events B, L mutually exclusive? If yes, what does it say for the pitcher B?
Are the events A, W independent? If yes, what does it say for the pitcher A?
Are the events B, L dependent? If yes, what does it say for the pitcher B?

The translation of the mathematics notation to the daily language plays an important role in understanding and solving problems related with probability. The following table shows the translation of these languages from one another.

Mathematics	Notation	$AW=\emptyset$
	Interpretation	The intersection of the two sets A and W is empty
Statistics		The sets A and W are mutually exclusive
English	Symbolic	"Winning the game" and "A pitching the game" can not occur at the same time
	Interpretation	If A pitches then we know that the team will not win the game. Therefore, A is a terrible pitcher.

Now, fill out the following table for the A and W being independent.

Mathematics	Notation	
	Interpretation	
Statistics		
English	Symbolic	
	Interpretation	

What are the numbers that are given in the statement of the problem?

First of all the problem says A has a winning average 0.4. So A will win 40% of the games. But for A to win the game it should play the game as a pitcher. In other words A will win the game given that he plays the game. Therefore,

$P(W \mid A) = 0.40$
$P(W \mid B) = 0.60$
$P(W \mid C) = 0.80$.

Since these pitchers pitch with frequency 2, 3, and 5 out of every 10 games, respectively;

$P(A) = 0.20$
$P(B) = 0.30$
$P(C) = 0.50$.

The first part of the problem is asking for $P(W)$. What is the difference between $P(W)$ and any one of the following probabilities $P(W \mid A)$, $P(W \mid B)$, and $P(W \mid C)$?

The second part of the problem is asking for P(A | W). What is the difference between P(W | A) = 0.40 and P(A | W)?

What do you think that P(A | W) + P(B | W) + P(C | W) will be equal to?

Given the information on P(A), P(B), P(C), P(W | A), P(W | B), and P(W | C), Do you expect that P(A | W) will be greater than P(C | W) ? Why or why not?

Note that the components that will affect the P(A | W) are (i) A should play, that is P(A) and (ii) the team should win, that is P(W | A).

Now, we will solve the problem by using the 4 techniques that we introduced above.

Solution I. To find out **P(W)**, one might use the following logical steps:
(i) In which cases the team will win the game. The team will win the game when A pitches, or B pitches, or C pitches. Therefore

"The team wins the game" = "A pitches and the team wins"
or "B pitches and the team wins"
or "C pitches and the team wins".

By using our notation;

$W = (A \cap W) \cup (B \cap W) \cup (C \cap W) = AW \cup BW \cup CW$

Are the events AW, BW, CW mutually exclusive? Justify your answer.

$P(W) = P(AW \cup BW \cup CW) = P(AW) + P(BW) + P(CW)$ (Why?)

Do we know the value of P(AW), P(BW), and P(CW)?

Can we find P(AW), P(BW), and P(CW) given the information on P(A), P(B), P(C), P(W | A), P(W | B), and P(W | C)?

By using the multiplication rule (definition of the conditional probability), we end up with

$P(AW) = P(W \mid A)\, P(A) = 0.40 \times 0.20 = 0.08$
$P(BW) = P(W \mid B)\, P(B) = 0.60 \times 0.30 = 0.18$
$P(CW) = P(W \mid C)\, P(C) = 0.80 \times 0.50 = 0.40$

Note that the multiplication rule on P(AW) can be used also as follows:

$P(AW) = P(A \mid W)P(W)$. Why we did not use this form?

Is $P(AW) = P(A \mid W)P(W) = P(W \mid A)\, P(A)$? Justify your answer.

Therefore, $P(W) = P(AW) + P(BW) + P(CW)$
$= P(W \mid A)P(A) + P(W \mid B)\, P(B) + P(W \mid C)\, P(C)$
$= 0.08 + 0.18 + 0.40 = 0.66$

$$P(W) = P(W \mid A)P(A) + P(W \mid B)\, P(B) + P(W \mid C)\, P(C)$$

The second part of the problem is asking for **P(A | W)**. By using the definition of the conditional probability, we end up with

$$P(A \mid W) = \frac{P(AW)}{P(W)} = \frac{P(W \mid A)P(A)}{P(W \mid A)P(A) + P(W \mid B)P(B) + P(W \mid C)P(C)}$$

How did we get the second part of the above formula?

Therefore,

$P(A \mid W) = \frac{0.08}{0.66} = 0.121$. Similarly $P(B \mid W) = 0.273$, and $P(C \mid W) = 0.606$.

We are going to interpret these results after we discuss all of the solution techniques.

Solution II. Tables. The information that is given in the problem can be represented by using a table. In this table the rows will show the result of the game, and the columns will show who pitches the game. The elements in the table will provide the probability that the corresponding row and column elements occur together.

	A PITCHES THE GAME (A)	B PITCHES THE GAME (B)	C PITCHES THE GAME (C)	
THE TEAM WINS THE GAME (W)	P(AW)	P(BW)	P(CW)	P(W)
THE TEAM LOSES THE GAME (W^c)	$P(AW^c)$	$P(BW^c)$	$P(CW^c)$	$P(W^c)$
	P(A)	P(B)	P(C)	1

Note that this table gives us some important formulas such as

$$P(AW) + P(BW) + P(CW) = P(W)$$

$$P(AW^c) + P(AW^c) + P(AW^c) = P(W^c)$$

If the conditional probabilities are available table can be modified by using the multiplication rule as follows;

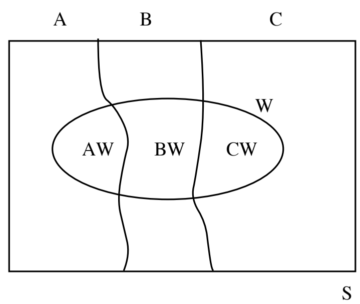
	A PITCHES THE GAME (A)	B PITCHES THE GAME (B)	C PITCHES THE GAME (C)	
THE TEAM WINS THE GAME (W)	P(AW)=P(W|A)P(A)	P(BW)=P(W|B)P(B)	P(CW)=P(W|C)P(C)	P(W)
THE TEAM LOSES THE GAME (W^c)	$P(AW^c) = P(W^C \mid A)P(A)$	$P(BW^c) = P(W^C \mid B)P(B)$	$P(CW^c) = P(W^C \mid C)P(C)$	$P(W^c)$
	P(A)	P(B)	P(C)	1

Now, let us fill out the table with the probabilities that are given

	A PITCHES THE GAME (A)	B PITCHES THE GAME (B)	C PITCHES THE GAME (C)	
THE TEAM WINS THE GAME (W)	$0.4 \times 0.2 = 0.08$	$0.6 \times 0.3 = 0.18$	$0.8 \times 0.5 = 0.40$	P(W)=?=
THE TEAM LOSES THE GAME (W^c)	$P(W^C \mid A)P(A)$	$P(W^C \mid B)P(B)$	$P(W^C \mid C)P(C)$	$P(W^c)$=?=
	0.20	0.30	0.50	1

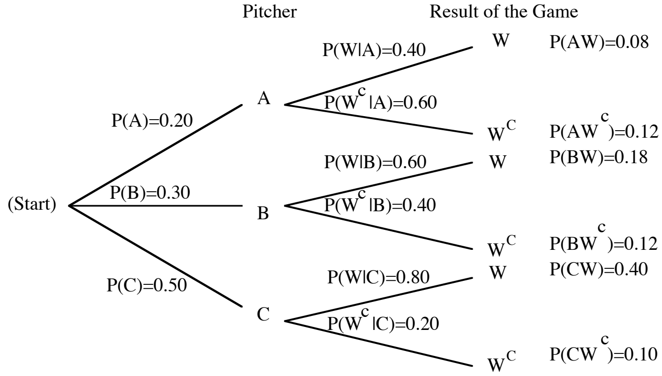
Fill out the rest of the table, and find P(W) and P(A | W).

Solution III. This time we will use a Venn diagram. Each region on the diagram will correspond to a relevant event.



Since the probability is a special type of measure, just like the area of each piece on the graph, same type of argument that is given in the solution I can be used to answer to the question.

Solution IV. Trees. When we analyze an experiment that takes place in a sequence of steps, we often find it convenient to represent the sample space by the set of all paths through the tree. A probability measure is assigned by considering the conditional probabilities appropriate for the outcome of the any step, given all previous outcomes. These weights are assigned at the appropriate branches of the tree, and then the weights for a path through the tree is the product of the branch weights along the path. In our experiment first of all one of the pitchers should pitch the game then the team either will win or lose. Therefore corresponding tree can be constructed as follows:



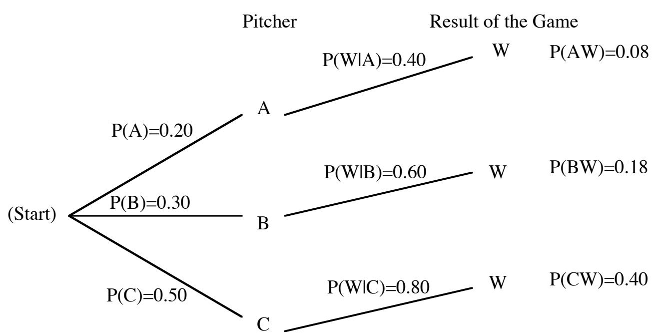
Mark the paths on the tree where the team ends up winning the game. By using these paths write down the formula for finding P(W)

P(W) =

Find the probability that the team will win the game by using the above formula

P(W) =

Suppose that the team wins the game. Given this information some of the branches of the tree will be chopped off.



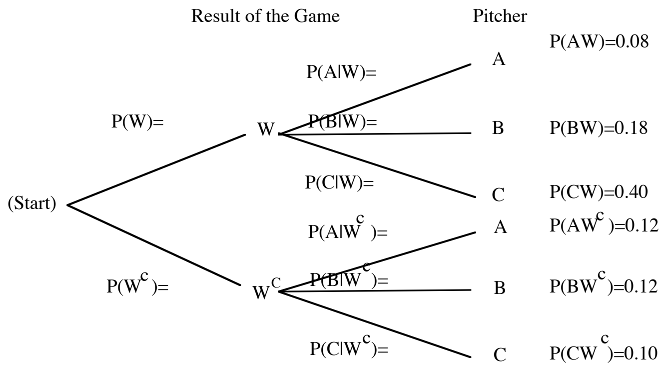
Write down the formula to find P(A | W)

P(A | W) =

Use the formula to find P(A | W)

P(A | W) =

Our original tree gave us the probabilities for the result of the game given the pitcher. The *i.inverse probabilities*, that is probabilities for the pitcher given the result of the game are called *Bayes probabilities*. These inverse probabilities can also be obtained by simply constructing the three for the two-stage experiment carried out in reverse order.



Fill out the rest of the above reverse tree by using P(W) and P(W^c) that you have already found.

Interpretation of the Results

In this part we will try to interpret the results that we got. Note that if the pitcher C pitched all the games the probability of winning would be 0.80. The probability of winning reduces to 0.66 since all the games are not pitched by C and all the pitcher are not equally good.

If all the pitchers are equally good, say with the winning percentage 0.80, what would be the probability that the team will win?

The winning percentages for the pitchers A, B, and C are 0.40, 0.60, and 0.80 respectively. And A, B, C pitches 20%, 30%, and 50% of the games.

Given the winning percentages for each pitcher, is it reasonable if A, B, and C pitches 50%, 30%, and 20% of the games?

If the team is winning the probability that A is pitching the game is 0.121. If all the pitchers were equally good, that is winning is independent to the whether A is pitching or not, what would this probability be equal to?

What type of relation do you expect between P(C | W) and P(C)? (P(C | W) > P(C) or
P(C | W) < P(C) or P(C | W) = P(C))

Case Study Revisited

Now we are ready to answer the questions that we rose in our case study. The table of frequencies can be changed into empirical probabilities as follows. For example let us try to find the probability of selecting a case which involves white defendant, white victim, and a death penalty verdict. There are 19 cases out of 326 in which all of these events occurred together. Therefore the chance of observing this event is 19/326=0.058.

Define the events of interest for the case study.

A = "receiving a death penalty verdict"
B = A^c =
C = "white defendant"
D = C^c =
E = "white victim"
F = E^c =

What is the meaning of A and C being mutually exclusive?

What is the meaning of A and C being independent?

What is the meaning of C and E^c being mutually exclusive?

What is the meaning of C and E being independent?

If there is no discrimination, what type of relation you expect to see between A and C? That is, do they have to be mutually exclusive, or independent, or both?

Which one of the following probabilities will give you a better information on discrimination? P(A | C) or P(C | A). Similarly P(A | D) or P(D | A).

Construct the marginal table for defendant's race and the death penalty verdict by ignoring victim's race.

	DEATH PENALTY		
DEFENDANT'S RACE	Yes	No	Total
White			
Black			
Total			

Find

P(A)

P(A | C) =

P(A | D) =

P(C | A) =

P(D | A) =

Is there an evidence of racial discrimination?

Construct the same table by controlling for the victim's race. That is create a table for defendant's race and death penalty verdict for white victims and black victims.

For white victims

DEATH PENALTY			
DEFENDANT'S RACE	Yes	No	Total
White			
Black			
Total			

For black victims

DEATH PENALTY			
DEFENDANT'S RACE	Yes	No	Total
White			
Black			
Total			

Find

P(A | CE) = P("death penalty" | "white victim" and "white defendant') =

P(A | DE) =

P(A | CF) =

P(A | DF) =

Is there an evidence of racial discrimination when we control for the race of the victim?

Construct a marginal table for the defendant's race and victim's race.

DEATH PENALTY			
DEFENDANT'S RACE	Yes	No	Total
White			
Black			
Total			

Find

P(E | C) =

P(E | D) =

P(F | C) =

P(F | D) =

What type of relation we have between defendant's race and victim's race?

I.6. DRUNKARD'S WALK:

RANDOM WALK:
WHERE IS A DRUNK PERSON GOING TO END UP?
MOTIVATION: SCENARIO

Most of the problems we encounter in daily living may be modeled probabilistically. The probability models can help us to understand the nature of the problems that we are facing better. Elementary probability models assume that the possible outcomes for each experiment are the same and the outcome of one experiment does not influence the outcomes of the other experiments probabilistically. In some cases these assumptions are not valid. For example;

* how much money you are going to have on your saving account next year depends on how much money you have on this account this year,
* next year's interest and unemployment rates will be affected by the this year's rates,
* a gambler's fortune at a game depends on the gambler's fortune at the previous game (Gambler's Ruin Problem)
* the population size of an endangered species in year 2000 depends on the population of this species in previous years.

BIG QUESTION

Which leads to the big question: How do we model a phenomenon where knowledge of the outcome of previous experiments influences our predictions for the outcomes of the next experiment?

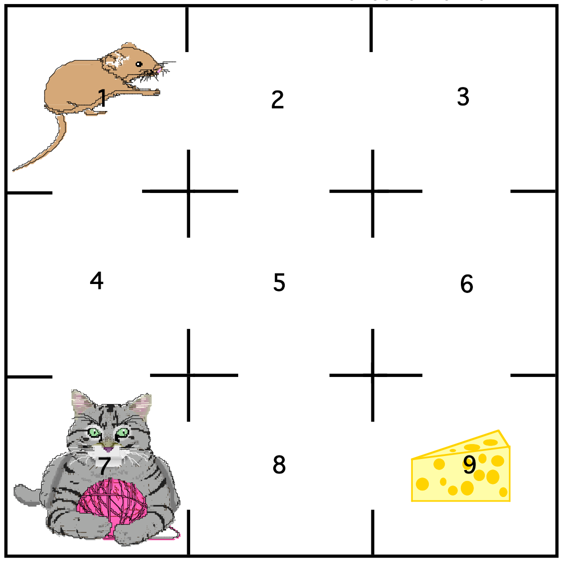
LEARNING OBJECTIVES

Learn how to model chance processes for which the knowledge of previous outcomes influences prediction for future experiment. After completing this activity you should be able to construct a probability model for the cases where outcome in the future given the outcomes in the past and present depends only on the outcome that we observe in the present. You should also be able to see how different tools, such as tree diagrams, graphs, matrices, and tables, can be used to simplify the problem.

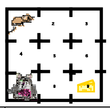
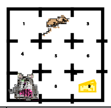
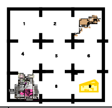
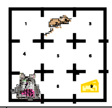
BACKGROUND: REVIEW, DEFINITIONS, FACTS & CONCEPTS

We will be learning Markov chains that are a type of a *stochastic process*. *A stochastic process* is a squenceof random variables. One can represent a stochastic process as $\{X(t), t\in \mathbf{T}\}$ where for each $t\in\mathbf{T}$, $X(t)$ is a random variable. The index t is often interpreted as time and, as a result, we refer to $X(t)$ as the state of the process at time t. **T** is called the *index set* of the process. If the index set is countable, the stochastic process is a *discrete-time process*. If **T** is an interval of the real line, the stochastic process is said to be a *continuous-time process*. Generally, the discrete-time process is represented by X_n instead of $X(t)$. **S**, *state space*, is the set of all possible random variables, $X(t)$, of the *stochastic process*. A realization of a stochastic process is called a *sample path* of the process.

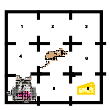
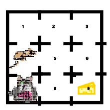
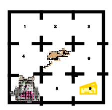
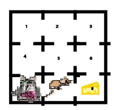
As *warm-up Activity*, consider the *Mouse in a Maze*. A hungry mouse starts in cell 1, while a cat is hiding patiently in cell 7, and the cheese is in cell 9. The mouse moves from cell to cell. . In the absence of learning, it will choose the next cell to visit randomly from the adjoining cells. Assume that once the mouse finds the piece of cheese or the cat, it will understandably stay there forever.



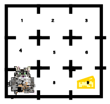
Here is a sample path for this example:

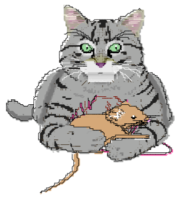
Step (n)	0	1	2	3
X_n	1	2	3	2

Step (n)	4	5	6	7
X_n	5	4	5	8



Step (n)	8
X_n	7



What do you understand when one says
- $X_0=1$
- $X_5=4$
- $X_8=7$

What is the state space **S** of this process?

First of all, why is this a *stochastic process?*

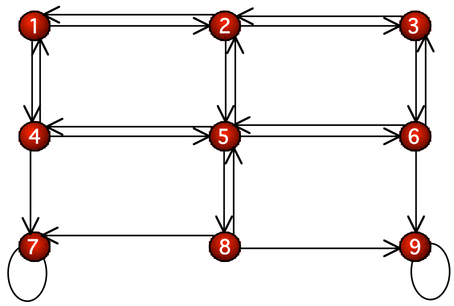
Study the following matrix; it is called a *one-step transition probability* matrix. What ideas do you have about how this matrix was constructed and what it means?

$$P = \begin{bmatrix} 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 1/3 & 0 & 1/3 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 0 & 1/3 & 0 & 1/3 & 0 & 0 \\ 0 & 1/4 & 0 & 1/4 & 0 & 1/4 & 0 & 1/4 & 0 \\ 0 & 0 & 1/3 & 0 & 1/3 & 0 & 0 & 0 & 1/3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/3 & 0 & 1/3 & 0 & 1/3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Study the following matrix; it is called an *incidence matrix*. What ideas do you have about how this matrix was constructed and what it means?

$$I = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Study the following graph; it is called a *state diagram*. What ideas do you have about how this graph was constructed and what it means?



The following matrix is obtained by multiplying the matrix I by itself. If you do not understand how to multiply two matrices do not worry. Just tell us what you think that this matrix is representing.

The following matrix is obtained by multiplying the matrix I by itself three times. What do you think that this matrix representing?

$$I^3 = \begin{bmatrix} 0 & 5 & 0 & 4 & 0 & 3 & 1 & 2 & 0 \\ 5 & 0 & 5 & 0 & 8 & 0 & 3 & 0 & 3 \\ 0 & 5 & 0 & 3 & 0 & 4 & 0 & 2 & 1 \\ 4 & 0 & 3 & 0 & 6 & 0 & 4 & 0 & 2 \\ 0 & 8 & 0 & 6 & 0 & 6 & 2 & 4 & 2 \\ 3 & 0 & 4 & 0 & 6 & 0 & 2 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 2 & 0 & 2 & 0 & 4 & 0 & 3 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

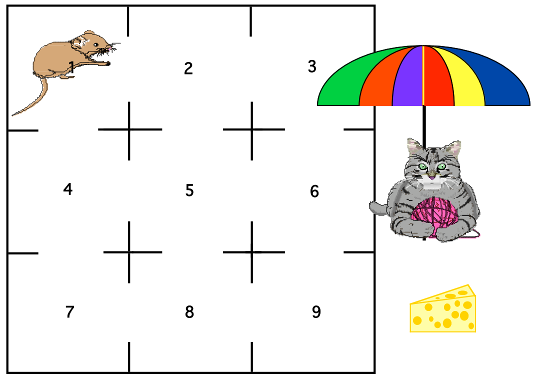
The following matrix is obtained by multiplying the matrix *I* by itself four times. What do you think that this matrix representing?

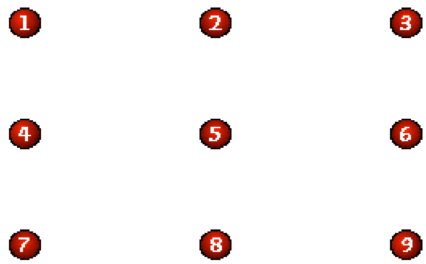
$$I^4 = \begin{bmatrix} 9 & 0 & 8 & 0 & 14 & 0 & 7 & 0 & 5 \\ 0 & 18 & 0 & 13 & 0 & 13 & 3 & 8 & 3 \\ 8 & 0 & 9 & 0 & 14 & 0 & 5 & 0 & 7 \\ 0 & 13 & 0 & 10 & 0 & 9 & 4 & 6 & 2 \\ 14 & 0 & 14 & 0 & 24 & 0 & 12 & 0 & 12 \\ 0 & 13 & 0 & 9 & 0 & 10 & 2 & 6 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 8 & 0 & 6 & 0 & 6 & 3 & 4 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Suppose that the cat is on vacation and you are given the following incidence matrix.

$$I = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

Draw the maze (some of the doors between the cells are now closed)
Tell us where the cheese is.



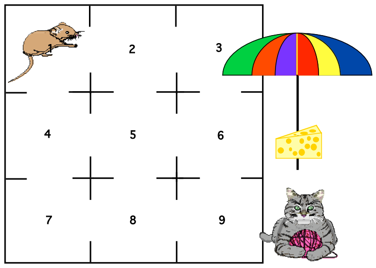


Complete the following state diagram.

Do you think that drawing the above graph would help you to answer the last two questions?

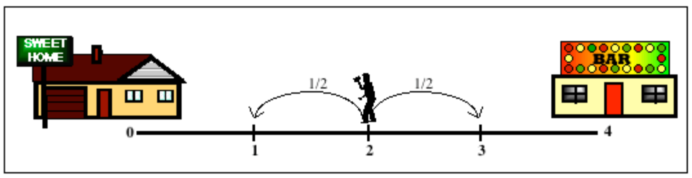
Now, let us challenge ourselves. (I do not have slightest idea at this point whether this can be solved or not) This time it is the Cheese's turn to take a vacation. You are given the following two-step incidence matrix.

Draw the maze.
Find where the cat is.



LEARNING ACTIVITY

Suppose that a person is walking along a four-block stretch of Park Avenue. The person starts at corner x and with probability 1/2, walks one block to the right and, with probability 1/2, walks one block to the left, when the person comes to the next corner she/he again randomly chooses her/his direction. The person continues until she/he reaches corner 4, which is a bar, or corner 0, which is a home. If the person reaches either home or the bar, she/he stays there. Here is a graphical representation of the problem.



1. You will need a coin or other method of simulating an event with a probability 1/2. Pick a starting point 1, 2, or 3. Toss a coin. If the coin shows head move to the right, if it is tail move to the left. Continue tossing the coin until you reach to the bar or the home.

For example, if you get the following sequence and start at 1: H H T H T
Your position at each step will be: 2 3 2 3 2.
Since you could not reach the bar of the home you should continue tossing the coin.

2. a. Starting from 1, repeat the above experiment 10 times, fill out the following table, and record the number of times you ended up home and bar.

COIN TOSS SEQUENCE	**POSITION**	**DESTINATION**
H H T H T T T	2 3 2 3 2 1 0	Home
TOTAL NUMBER OF HOMES		

b. Starting from 2, repeat the above experiment 10 times, fill out the following table, and record the number of times you ended up home and bar.

COIN TOSS SEQUENCE	POSITION	DESTINATION
H H T H T T T	3444444	Bar
TOTAL NUMBER OF HOMES		

c. Estimate you chance of reaching home starting from 1 by using the results of the experiments? What about your chance of reaching the bar?

Estimate you chance of reaching home starting from 2 by using the results of the experiments? What about your chance of reaching the bar?

Compare your results and justify why these estimates are different.

CONNECTIONS

3. ***TREE DIAGRAMS***

a. Consider the following tree which represents all possible outcomes of tossing a coin 4 times. Follow the individual branches and determine where the drunk person will be if she/he starts at 3.

b. Since, if the drunk person reaches either home or the bar, she/he stays there, eliminate the unnecessary branches of the tree diagram given above.

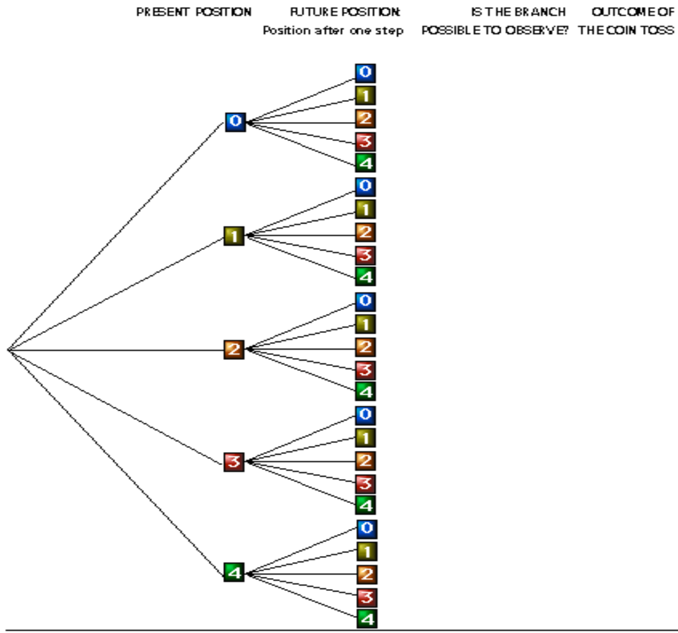
c. Is it possible that starting from block 3 this drunk person will never be able to reach home or the bar? Give an example.

4. FROM TREES TO MATRICES

As you can see representing this problem in a compact form by using the trees is difficult. You have to be able to produce tree diagrams with infinite number of branches for each starting point. Now let us search for a better way of summarizing this problem.

a. Suppose that we saw this person at one of the blocks, the person had enough time to take only one step. Where this person will be when we come back? What are the factors that will affect your answer?

b.



c. Fill out the following table cross out the impossible steps, for the remaining ones put down the corresponding outcome of the coin toss experiment.

	Future Position (Position after one step)				
Present Position	0	1	2	3	4
0					
1					
2					
3					
4					

Replace corresponding outcomes of the experiment with the probability that matches. Now, we can represent the table of probabilities in a compact form as follows:

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Interpret this ***matrix*** by referring to the table in c. This matrix is called ***one-step transition matrix.*** The states 0 and 4 are impossible to leave such states are called ***absorbing states.*** States which are not absorbing called ***transient.***

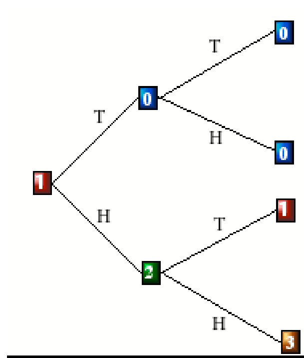
5. *MARKOV PROPERTY*

Suppose that the drunk person was in block 1, then 2, then 1, then 2 and now the person is in block 3. You want to know where the person will be after the next step. Do you need to know that the person was in corners 1, 2, 1, 2, to determine where the person will be after visiting block 3?

Suppose that the present the person is in block 3, does it make any difference if this is the persons first, second, or tenth visit of block 3, to determine where the person will be after the next step?

6. *WHAT WILL HAPPEN AFTER TWO OR MORE STEPS?*

Suppose that the drunk person was in block 1, where this person could be after two steps? Construct a tree diagram to show possibilities. As an example we have given below a tree diagram when the drunk person was in block 1.



Fill out the following table cross out the impossible steps, for the remaining ones put down the corresponding outcome of the coin toss experiment. Note that this time we have to toss the coin two times.

	Position after ***two*** step				
Present Position	0	1	2	3	4
0					
1					
2	TT	-	TH or HT	-	HH
3					
4					

Replace corresponding outcomes of the experiment with the probability that matches. Now, we can represent the table of probabilities in a compact form as follows:

$$P^{(2)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1/2 & 1/4 & 0 & 1/4 & 0 \\ 1/4 & 0 & 1/2 & 0 & 1/4 \\ 0 & 1/4 & 0 & 1/4 & 1/2 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Interpret this ***matrix*** which is called ***two-step transition matrix.*** Verify that

$$P^{(2)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

In fact if you want to know what will happen probabilistically after n steps, n-step transition matrix will be:

$$P^{(n)} = P^n = P \times P \times \cdots \times P$$

By using a calculator or MINITAB find the 4-step transition matrix and interpret.

```
MTB > read c1-c5
 DATA> 1 0 0 0 0
 DATA> 0.5 0 0.5 0 0
 DATA> 0 0.5 0 0.5 0
 DATA> 0 0 0.5 0 0.5
 DATA> 0 0 0 0 1
 DATA> end
      5 ROWS READ
 MTB > copy c1-c5 m1
 MTB > print m1
  MATRIX M1
   1.0  0.0  0.0  0.0  0.0
   0.5  0.0  0.5  0.0  0.0
   0.0  0.5  0.0  0.5  0.0
   0.0  0.0  0.5  0.0  0.5
   0.0  0.0  0.0  0.0  1.0
 MTB > multiply m1 m1 m2
 MTB > print m2
  MATRIX M2
   1.00  0.00  0.00  0.00  0.00
   0.50  0.25  0.00  0.25  0.00
   0.25  0.00  0.50  0.00  0.25
   0.00  0.25  0.00  0.25  0.50
   0.00  0.00  0.00  0.00  1.00
```

```
MTB > multiply m1 m2 m3
 MTB > print m3
  MATRIX M3
   1.000  0.000  0.000  0.000  0.000
   0.625  0.000  0.250  0.000  0.125
   0.250  0.250  0.000  0.250  0.250
   0.125  0.000  0.250  0.000  0.625
   0.000  0.000  0.000  0.000  1.000
MTB > multiply m1 m3 m4
 MTB > print m4
  MATRIX M4
   1.000  0.000  0.000  0.000  0.000
   0.625  0.125  0.000  0.125  0.125
   0.375  0.000  0.250  0.000  0.375
   0.125  0.125  0.000  0.125  0.625
   0.000  0.000  0.000  0.000  1.000
```

ONGOING ASSESSMENT AND WRAP-UP

7. Discuss the implications of Markov Properties. List down possible tools that can be used to model a problem. Which one of these tools is the most powerful in your opinion? Why?

8. Give some examples where this process can be used in the real world.

EXTENSIONS

9. **(Forecasting the weather)** The chance of rain tomorrow depend on only today's whether. If it rains today, then it will rain tomorrow with probability 0.7 and if it does not rain today then it will rain tomorrow with probability 0.4. Form the one-step transition matrix.

Find the two-step transition probability matrix. What is the probability that it will rain two days from today given that it is remaining today?

Find the four-step transition probability matrix and interpret. Did you notice any interesting pattern in probabilities?

ANSWERS:

$$P = \begin{matrix} Rain \\ No\ Rain \end{matrix} \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix}$$

$$P^{(2)} = \begin{matrix} Rain \\ No\ Rain \end{matrix} \begin{bmatrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{bmatrix}$$

$$P^{(4)} = \begin{matrix} Rain \\ No\ Rain \end{matrix} \begin{bmatrix} 0.5749 & 0.4251 \\ 0.5668 & 0.4332 \end{bmatrix}$$

10. **(War of long distance companies)** It has been observed that customers switch from one long distance company to another according to the following transition matrix.

$$P = \begin{matrix} MCI \\ AT\&T \\ SPRINT \end{matrix} \begin{bmatrix} 3/4 & 1/4 & 0 \\ 0 & 2/3 & 1/3 \\ 1/4 & 1/4 & 1/2 \end{bmatrix}$$

Find the two-step transition matrix and interpret one and two-step transition matrices. Find the probability that a customer who uses MCI today will again be using MCI two months from today, assuming that she or he switches once a month.

$$P^{(2)} = \begin{matrix} MCI \\ AT\&T \\ SPRINT \end{matrix} \begin{bmatrix} 9/16 & 17/48 & 1/12 \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{bmatrix}$$

Find the twelve step transition matrix. Interpret the matrix.

ANSWER:

$$P^{(\infty)} = \begin{matrix} MCI \\ AT\&T \\ SPRINT \end{matrix} \begin{bmatrix} 2/7 & 3/7 & 2/7 \\ 2/7 & 3/7 & 2/7 \\ 2/7 & 3/7 & 2/7 \end{bmatrix}$$

11. **(Gossip spreading)** If a member of a community is entrusted with some news, she/he will tell it to her/his friends truly and without any changes with probability p. The probability that she/he will twist it around and tell the opposite is really negligible, say, $1-p$. The probability of such a misconduct and misbehavior is less than one in a billion. Therefore, if a member of the community receives news, she/he will tell it to another member truly and unchanged with probability *p*, and she/he will twist it and tell its opposite with probability $1-p$. Let

T= "the true statement about the young person's dog"

N= "the negation of the true statement about the person's dog".

Construct the one-step transition matrix. Interpret the matrix. Find ten, twenty, one hundred-step transition probability matrices. Do you see anything interesting? How are the probabilities changing? Can you say that in general in this community only half of the people is telling the truth?

$$P = \begin{matrix} T \\ N \end{matrix} \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}$$

$$P^{(\infty)} = \begin{matrix} T \\ N \end{matrix} \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$$

12. Visit the Mouse in a Maze example and place the cat so that the chance that the mouse will end up with cheese will be lower than the cat catching the mouse.

I.7. A Markov Chain Activity

MATH. 3610: INTRODUCTION TO PROBABILITY THEORY & STOCHASTIC PROCESSES

MARKOV CHAINS

Objective: To learn how to set up probability transition matrix and construct an appropriate state diagram.

Activity:

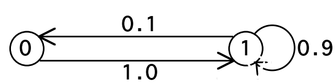
1. Consider a farmer using an old tractor. The tractor is often in the repair shop but it always takes only one day to get it running again. The first day out of the shop it always works but on any given day thereafter, independently of its previous history, there is a 10% chance of it not working and thus being sent back to the shop.

Let X_0, X_1, ... be random variables denoting the daily condition of the tractor, where a one denotes the working condition and a zero denotes the failed condition.

Explain in words
"X_n=1"→
"X_n=0"→
Write one step transition probability matrix

P=

Here is the state diagram for this Markov chain:



2. A salesperson lives in town *a* and is responsible for towns *a, b, and c*. Each week he is required to visit a different town. When s/he is in her/his home town, it makes no difference which town s/he visits next so s/he flips a coin and if it heads s/he goes to *b* and if tails s/he goes to *c*. However, after spending a week away from home s/he has a slight preference for going home so when s/he is in ether towns *b* or *c* s/he flips two coins. If two heads occur, then s/he goes to the other town; otherwise s/he goes to *a*.

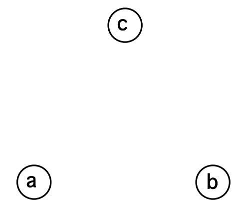
Define
X_0=
X_1=
X_{16}=

Explain in words
"X_n=a"→
"X_n=b"→
"X_n=c"→

Write one step transition probability matrix

P=

Construct a state diagram for this Markov chain:



3. Let $\{X_n;\ n=0, 1, ...\}$ be a Markov chain with state space {1, 2, 3, 4} and transition probabilities given by

$$\mathbf{P}= \begin{Vmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.3 & 0.7 & 0 \\ 0 & 0.5 & 0.5 & 0 \\ 0.2 & 0 & 0.1 & 0.7 \end{Vmatrix}$$

Construct a state diagram for this Markov chain:

Starting from state 4 is it possible not to come back state 4 again? Give an example.
What happens if you start with state 1.

4. This example illustrates that the *n* need not refer to time. Consider a page of text and represent vowels by zeroes and consonants by ones. Thus the page becomes a string of zeros and ones. It has been indicated that the sequence of vowels and consonants in the Samoan language forms a Markov chain, where a vowel always follows a consonant and a vowel follows another vowel with a probability of 0.51.
Write one step transition probability matrix

P=

ASSESSMENT

Each morning an individual leaves her house and goes for a run. She is equally likely to leave either from her front or back door. Upon leaving the house, she chooses a pair of running shoes (or goes running

barefoot if there are no shoes at the door from which she departed). On her return she is equally likely to enter, and leave her running shoes, either by the front or back door. Suppose that she owns a total of 5 pairs of running shoes.
Let X_n denote the number of pairs of shoes at the door the runner departs from at the beginning of day n.

Write one step transition probability matrix

P=