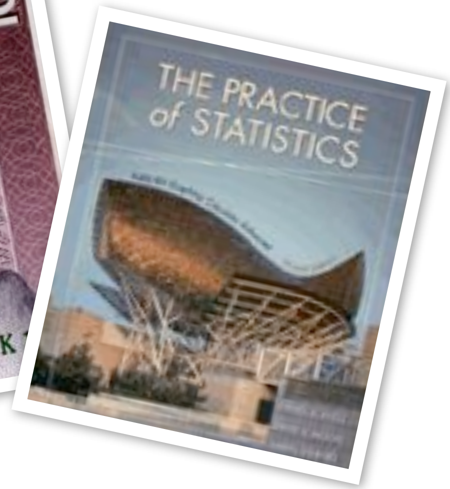
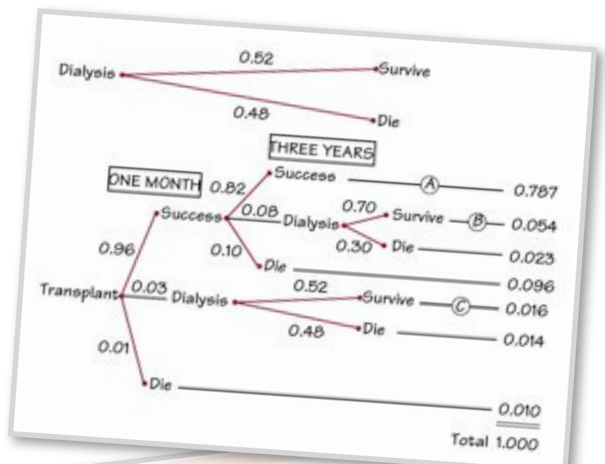
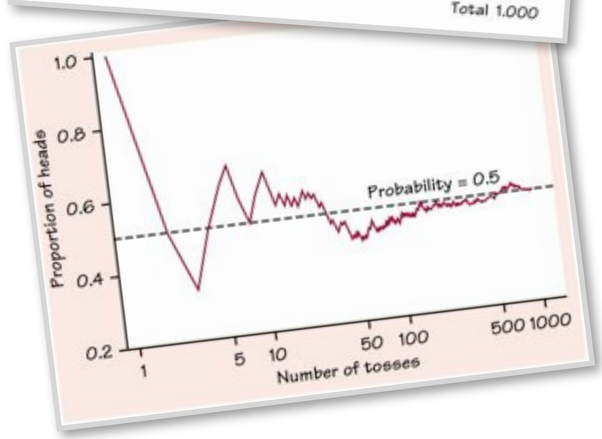


Chapter 6

Probability: The Study of Randomness

Chance is all around us. Probability is the branch of mathematics that describes the pattern of chance outcomes. It is important that we understand chance outcomes since the reasoning of statistical inference rests on asking, "How often would this outcome occur by chance alone?"

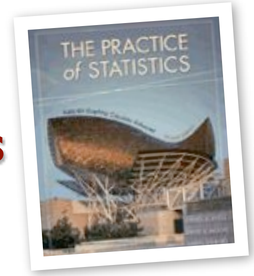




Probability:

- 6.1: The Idea of Probability
- 6.2: Probability Models
- 6.3: Probability Rules

AP Statistics Chapter 6: Probability: The Study of Randomness



"The most important questions of life are, for the most part, really questions of probability" ~LaPlace

Tentative Lesson Guide					
Date		**Stats**	**Lesson**	**Assignment**	**Done**
Mon	11/27	6.1	Flipping Coins…Randomness	Rd 330-340 Do **11-15, 17-18**	
Tues	11/28	6.2	Probability Rules	Rd 342-354 Do **19-23, 26-29, 31**	
Wed	11/29	6.3	General Probability Rules	Rd 359-369 Do-**53 46**	
Thu	11/30	6.3	Conditional Probabilities	Rd 366-369 Do **54-61**	
Fri	12/1	**Quiz**	**Quiz - Probability**		
Mon	12/4	6.3	Bayes' Rule - Practice	Rd 371-376 Do **62-65**	
Tues	12/5	**Rev**	**Review**	Rd 383-383 Do **66-71, 73-74**	
Wed	12/6	*2 Hour Late Start - Opportunities Day*			
Thurs	12/7	**Rev**	**Review**	Do **78-87**	
Fri	12/8	**Exam**	**Exam Chapter 6**	**Online Quiz Due**	

Note:
The purpose of this guide is to help you organize your studies for this chapter.
The schedule and assignments may change slightly.

Keep your homework organized and refer to this when you turn in your assignments at the end of the chapter.

Class Website:
Be sure to log on to the class website for notes, worksheets, links to our text companion site, etc.

http://web.mac.com/statsmonkey

Don't forget to take your online quiz!. Be sure to enter my email address correctly!
http://bcs.whfreeman.com/yates2e

My email address is:

jmmolesky@isd194.k12.mn.us

Chapter 6 Objectives and Skills:

These are the expectations for this chapter. You should be able to answer these questions and perform these tasks accurately and thoroughly. Although this is not an exhaustive review sheet, it gives a good idea of the "big picture" skills that you should have after completing this chapter. The more thoroughly and accurately you can complete these tasks, the better your preparation.

The Idea of Probability

Random phenomenon...individual outcomes are uncertain, but there is a regular distribution of outcomes in a large number of repetitions.

Probability...proportion of times an event occurs in many repeated events of a random phenomenon. Probability can also be thought of as long-term relative frequency.

Independent events (trails)...outcome of an event (trial) does not influence the outcome of any other event (trial).

Probability Models

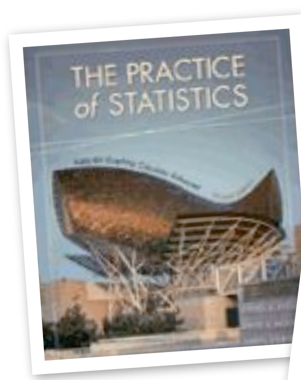
- [ ] Sample space: The set of all possible outcomes of a random phenomenon
- [ ] Event: any outcome or a set of outcomes of a random phenomenon
- [ ] Probability model: a mathematical description of a random phenomenon consisting of two parts: a sample space "S" and a way of assigning probabilities to particular events.
- [ ] It is very important to find all possible outcomes in a sample space. To do this, you may want to use a tree diagram or the multiplication counting principle.
- [ ] If A is an event, then 0 ≤ P(A) ≤ 1
- [ ] Complement rule: P(not A) = 1 - P(A).
- [ ] Two events are disjoint, or mutually exclusive, if they have no outcomes in common. If A and B are disjoint, then P(A or B) = P(A) + P(B).
- [ ] Two events are independent if knowing that one occurs does not change the probability of the other. If events A and B are independent, then
  P(A and B) = P(A)P(B).

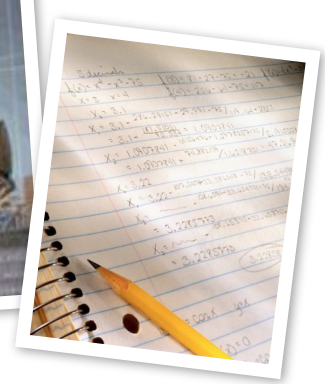
General Probability Rules

- [ ] There are basic laws that govern uses of probability. An understanding of these laws is important for those who wish to utilize statistics in practical and meaningful ways.
  - [x] For any two events A and B,
  - [x] P(A or B) = P(A) + P(B) - P(A and B).
  - [x] If A and B are disjoint, then P(AandB)= 0.
  - [x] Prob(B|A) = P(A and B)/P(A).
- [ ] The probability formulas are useful, but often times they are not needed if sample spaces are small and one uses common sense.
- [ ] In many problems, a tree diagram may help organize information and make calculating probabilities somewhat easier. One type of problem in which tree diagrams are especially useful is that of the "converse problem" as studied by Thomas Bayes. His findings, known as Bayes' Rule or Bayes' Theorem, allow us to solve problems such as "Given somebody tests positive for a disease, what is the probability they actually have that disease?"
- [x] **Bayes' Rule**

$$P(A \mid B) = \frac{P(A)P(B \mid A)}{P(A)P(B \mid A) + P(A')P(B \mid A')}$$

Rather than memorizing the formula, I'd suggest using a tree diagram and a little common sense...but that's just me.





6.1: Introduction to Probability - Flipping Coins

Adapted from Activity Based Statistics by Scheaffer, Gnanadeskian, Watkins, and Witmer

A central theme that runs throughout Tom Stoppard's play, "*Rosencrantz and Guildenstern are Dead,*" is the game of chance. Ros and Guil begin the play by flipping a coin only to discover that heads are produced consecutively. After the eighty-ninth flip, Guil begins to ponder this seeming anomaly in an attempt to explain how such a phenomenon could occur. "*List of possible explanations. One: I'm willing it. Inside where nothing shows, I am the essence of a man spinning double-headed coins, and betting against himself in private atonement for an unremembered past. Two: time has stopped dead, and the single experience of one coin being spun once has been repeated ninety-times. On the whole, doubtful. Three: divine intervention. Four: a spectacular vindication of the principle that each individual coin spun individually is as likely to come down heads as tails and therefore should cause no surprise each individual time it does.*"

Activity:

- ☑ Flip a coin 50 times and record the cumulative proportion of heads after each toss. To do this, use the Stirling recording sheet on the back of this handout. Consider heads a "win" and tails a "loss."
- ☑ When you have completed the 50 flips, record the final percentage of heads on the dotplot on the board. Also, enter your total number of heads on the board.

1) Describe the dotplot (sampling distributions) on the board. What does it suggest about the result of flipping a coin 50 times?

2) A key concept to remember when exploring randomness is that of *independence*. What does it mean if trials (successive flips) are said to be independent?

Probability Vocabulary:

☑ **Sample Space:**

☑ **Probability:**

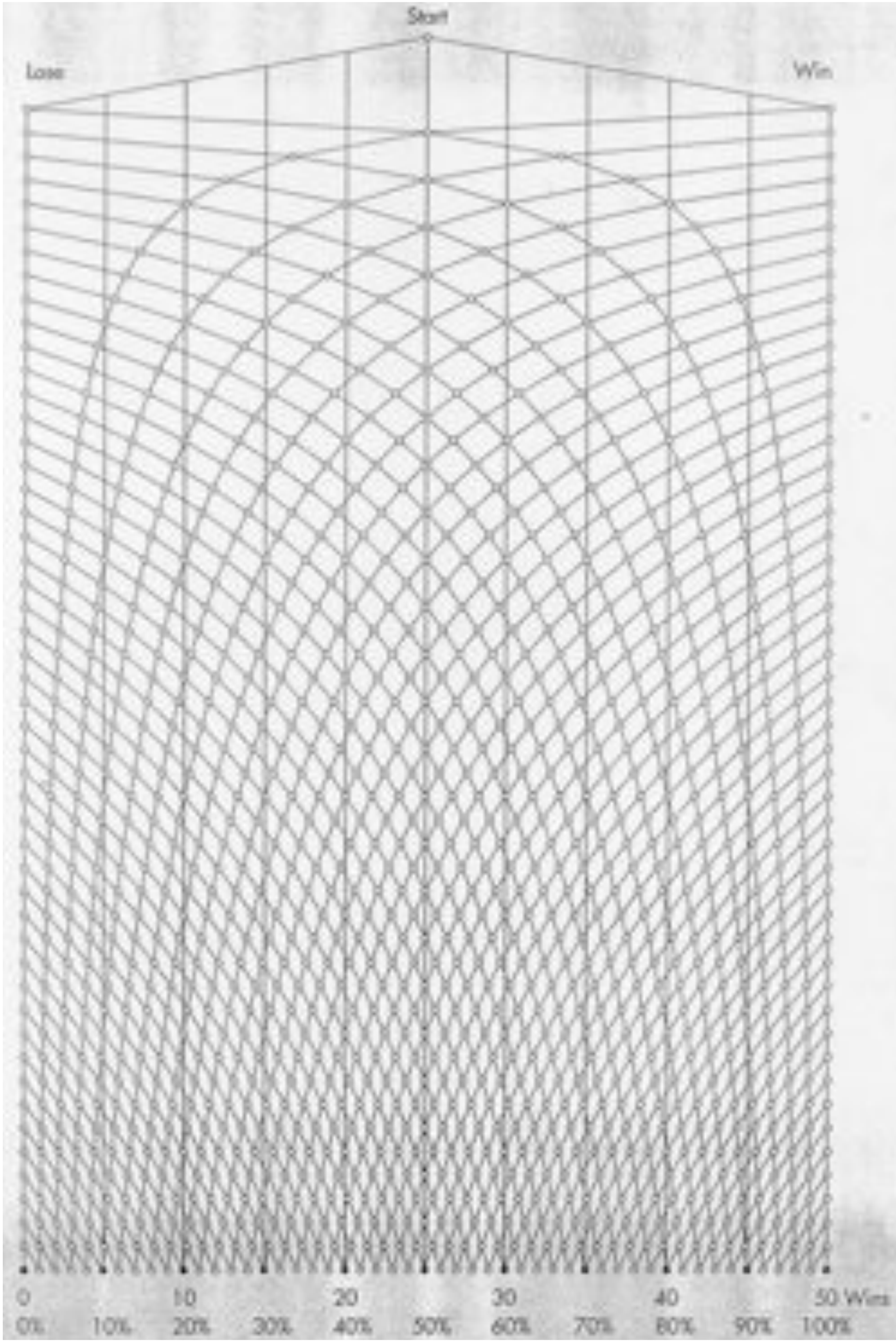
☑ **Probability Model:**

- ☑
- ☑

☑ **Independent:**

☑ **Random:**

☑ ***Key Concept*: Chance behavior is unpredictable in the short term, but has a regular and predictable pattern in the long run.**


Start
Lose
Win
0
10
20
30
40
50 Wins
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%

6.2: Probability Models

READ pages 333-354 in your text and complete the following notes:

The main idea in Chapter 6 is that random behavior is unpredictable in the short-term, but has a regular and predictable pattern in the long term. "**Random**" is a term used in statistics and probability theory to identify behavior that exhibits this unpredictability in single instances, but displays order and predictability in repeated trials.

The probability of an event can be calculated using a simple formula: **P(event)**=____________

To find P(event), one must know the possible outcomes of the situation--or at least the number of outcomes. A list of possible outcomes for any random phenomenon is called a _____________________.

We can find this list of outcomes in a number of different ways. One method is to branch out all of the possibilities to ensure all outcomes are represented. This is called a ______________________. Another method calculates the number of outcomes, but does not list the outcomes themselves. This is based on the _______________ _______________ ________________.

There are five basic rules of probability we need to be familiar with. Write the following rules in your own words:

- ☑ **$0 \leq P(A) \leq 1$**
- ☑ **P(S)=1**
- ☑ **$P(A^C)=1-P(A)$**
- ☑ **If AB=, then P(A or B) = P(A) + P(B)...otherwise, P(A or B) = P(A) + P(B)-P(A and B)**
- ☑ **If A and B are *independent*, then P(A then B) = P(A)P(B) for successive events.**

Example Probability Questions:

Consider the following probabilities defined by *Benford's Law*. This law states that the first digits of numbers in legitimate financial records follow a known distribution…namely:

Digit	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
P(digit)	**.301**	**.176**	**.125**	**.097**	**.079**	**.067**	**.058**	**.051**	**.046**

Let A={1st digit=1} B={1st digit≥6} C={1st digit is odd} D={1st digit<4}

Find P(A)=________ P(B)=__________ P(C)=_________ P(D)=___________

Verify Rule 2 holds:

P(B or D)= $P(D^C)$= P(C and D)=

P(B and C)= P(5 A's in a row)= P(C or D)=

6.3: Conditional Probability

One aspect of increased communication using the Internet is that diverse individuals can exchange opinions on various topics of interest. A side effect of such conversations is "*flaming*," that is, negative criticism of others' contributions. M. Dsilva studied this phenomenon in "Criticism on the Internet: An Analysis of Participant Reactions," *Communications Research Reports*, 1999. The investigators were interested in the effect that personal criticism has on an individual. Would being criticized make one more likely to criticize others? Data from the study are provided here:

	Have been flamed	Have not been flamed	Total
Have flamed others	**19**	**8**	**27**
Haven't flamed others	**23**	**143**	**166**
Total	**42**	**151**	**193**

Define the events F and O to be:

F = event that the individual has flamed others

O = event that the individual has been flamed by others

Find the following probabilities and describe what they tell us:

P(F) = P(F and O) = P(O) =

Suppose someone selected at random from this group admits they have been flamed. What is the probability they have flamed someone else? We can solve this by using a conditional probability. That is:
P(flamed someone given been flamed) = P(F | O)

<u>Conditional Probability {See Page369}</u>

The probability of B given A has occurred, denoted P(B | A), is

P(B | A) =

Find and interpret P(F | O) =

Find and interpret P(O | F) =

Are they the same? What can you conclude about flaming on the Internet?

<u>Handy Notes about Conditionals:</u>

- Most conditional probabilities can be figured out using common sense.
- Two events, A and B, are independent iff:
- General Multiplication Rule for Two or More Events: P(A and B) =

6.3: Conditional Probability Practice

1. If a single fair die is rolled, find P(3 | prime).

2. If two cards are drawn without replacement from a deck, find the probability that the second card is a diamond, given that the first card was a diamond.

3. If two fair dice are rolled, find the probability that the sum of the faces is 7, given that the first die rolled is odd.

4. If three cards are drawn without replacement from a deck, find the probability that the third card is a face card, given that the first card was a king and the second card was a 9.

5. A six-sided die is tossed. What is the probability that it shows 2 if you know the following?
   a) It shows an even number.
   b) It shows a number less than 5
   c) It does not show a 6.
   d) It shows 1 or 2
   e) It shows an even number less than 4.
   f) It shows a number greater than 3.

6. One container holds the letters D A D and a second container holds the letters A D D. One letter is chosen randomly from the first container and added to the second container. Then a letter will be chosen from the second container.
   a) What is the probability that the second letter chosen is D if the first letter was A? if the first letter was D?
   b) What is the probability that the second chosen letter is A if the first letter was A? if the first letter was d?

7. Let A and B be events with P(A)=1/3, P(B) =1/2 , and P(A and B)=1/6. Find P(A|B), P(B|A)

8. A couple wants to have 3 or 4 children, including exactly 2 girls. Is it more likely that they will get their wish with 3 children or with 4?

6.3: Bayes' Rule *Adapted from "Introduction to Statistics and Data Analysis" Peck, Olson, Devore*

Given sufficient information about the accuracy of a medical test for a disease, we can easily calculate the probability of the test giving a 'false positive' if *n* people are tested. However, is it possible to calculate the probability that a person has the disease given he or she tests positive? The Reverend Thomas Bayes (1702-1761) discovered a solution to what is known as the "converse problem." Consider the following information about a test for Lyme's disease from the article, "Laboratory Considerations in the Diagnosis and Management of Lyme Borreliosis" (Amer. J. of Clinical Pathology (1993)).

+	represents a positive result on the blood test
-	represents a negative result on the blood test
L	represents the event the patient actually has Lyme disease
L^c	represents the event the patient actually does not have Lyme disease

The following probabilities were reported in the article:

P(L)	= .00207	.207% of the population actually has Lyme disease
$P(L^c)$	= _______	_______% of the population does not have Lyme disease
P(+ | L)	= .937	93.7% of those with Lyme disease test positive
P(- | L)	= _______	_______% of those with Lyme disease test negative
$P(+ \mid L^c)$	= _______	_______% of those who do not have Lyme disease test positive
$P(- \mid L^c)$	= .97	97% of those who do not have Lyme disease test negative

Note, P(+ | L) represents the probability that a person tests positive given they have the disease. Bayes' converse problem poses the question: *Given a person tests positive, what is the probability they actually have the disease?* Bayes reasoned as follows to obtain the answer to the converse problem of finding P(L | +).

$$P(L \mid +) = \frac{P(\qquad)}{P(\quad)}$$

Note: Since _______________________, we can say ___________________.

Therefore...

$$P(L \mid +) = \frac{P(\qquad)P(\qquad\qquad)}{P(\quad)}$$

Bayes' Theorem:
If A and B are any events whose probabilities are not 0 or 1,

$$P(A \mid B) = \underline{\qquad\qquad\qquad\qquad\qquad\qquad}$$

6.3: Bayes' Rule Practice

1. Amazon.com allows you to choose from two shipping services (FedEx and UPS), both who offer overnight delivery of parcels. For those of us who do our Happy Festive Winter Season shopping at the last minute, it is extremely important that we get our packages by 10 AM. Suppose FedEx is used 70% of the time. UPS fails to meet the 10AM deadline 10% of the time, while FedEx delivers late 8% of the time.

Define the following events:

F = event that a customer chooses FedEx
U = event that a customer chooses UPS
L = event that the package was late.

Suppose a person received a package late. Which service is more likely to have been used?

2. In an article that appears on the web site of the ASA (www.amstat.org), Carlton Gunn, a public defender in Seattle, wrote about how he uses statistics in his work as an attorney. He states:

> I personally have use statistics in trying to challenge the reliability of drug testing results. Suppose the chance of a mistake in the taking and processing of a urine sample for a drug test is just 1 in 100. And your client has a "dirty" (ie. Positive) test result. Only a 1 in 100 chance that it can be wrong? Not necessarily. If the vast majority of all tests given – say 99 in 100 – are truly clean, then you get one false dirty and one true dirty in every 100 tests, so that half of the dirty tests are false.

Define the following events: TD=event that the test result is dirty, TC=event that the test result is clean, D=event that the person tested is using, C=event that the person tested is clean.

Using the information from the quote, find the following probabilities:

P(TD|D) = P(TD|C) = P(C) = P(D) =

Find P(TD) =

Use Bayes' rule to evaluate P(C|TD). Is this value consistent with Mr. Gunn's argument?

3. The article "Checks Halt over 200,000 Gun Sales" reported that required background checks blocked 204,000 sales in 1999. The article also indicated that state and local police reject a higher percentage than the FBI, stating, "The FBI performed 4.5 million of the 8.6 million checks, compared with 4.1 million by state and local agencies. The rejection rate among state and local agencies was 3%, compared with 1.8% for the FBI"

Use the given information to estimate P(F), P(S), P(R|F), and P(R|S). Use the probabilities to evaluate P(S|R) and interpret this value in the context of the problem.