Probability

Roulette is a nice, genteel way for a person to politely and quickly lose a great deal of money.

Jimmy “The Greek” Snyder

It is a truth very certain that, when it is not in our power to determine what is true, we ought to follow what is most probable.

René Descartes

IN THE REAL WORLD

1. Your favorite basketball team has the ball and trails by 2 points with little time remaining in the game. Should your team attempt a game-tying two-pointer or go for a buzzer-beating three-pointer to win the game? (This situation has often been used in Microsoft job interviews).
2. After a touchdown should a coach kick the extra point or go for two?
3. On 4th down should your favorite football team punt or try for the first down?
4. With a man on first base and no one out, should the manager call for a sacrifice bunt?
5. If your favorite basketball team has a 3 point lead with little time left on the clock and the other team has the ball, should your team foul?

Unit Objectives

At the conclusion of this unit you will be able to:

- 1) apply the rules of probability to calculate event probabilities in realistic situations
- 2) use probability trees to calculate probabilities when events are dependent.

Reading Assignment

Course text: Chapters 14, 15.

Highlights From The Readings

1) It is important to note that in a sense probability is the reverse of statistics: In probability we use the population information to infer the probable nature of the sample. In statistics we use the sample to make inferences about the population. Probability uses deductive reasoning; statistics uses inductive reasoning.

2) As the course progresses, you will see that many problems in sports, the environment, global health, and many other areas alternate between deduction and induction. For example, suppose a particular probability model for global warming is hypothesized; this model permits predictions of the future behavior of global temperature patterns. Temperature data is gathered and analyzed; the data indicate that the temperature patterns differ from what is predicted by our original probability model. We then adjust the original probability model based on the information from the sample data (induction). This deduction/induction process may be repeated several times in the course of attempting to understand and improve the global temperature process.
4.1 The Laws of Probability

NEWS CLIP
Irving Herzel, a professor from Iowa State used a computer to find the most probable squares on which you can land in the game of Monopoly.™
1. Illinois Avenue
2. Go
3. B.&O. Railroad
4. Free Parking

NEWS CLIP
Let the king prohibit gambling and betting in his kingdom, for these are vices that destroy the kingdoms of princes.
The Code of Manu, ca 100 A.D.

Historical Note
The mathematical theory of probability arose in France in the 17th century when a gambler, Chevalier de Méré, became interested in adjusting the stakes so that he could win more often than he lost. In 1654 he wrote Blaise Pascal, who in turn sent his questions to Pierre de Fermat. Together they developed the first theory of probability.

APPROACHES TO PROBABILITY

relative frequency approach

subjective probability approach

DEFINITIONS OF FUNDAMENTAL TERMS

Experiment:
act or process that leads to a single outcome that cannot be predicted with certainty

EXEMPLARY
1. Toss a coin
2. Attempt a 45-yard field goal
3. Bunt with a runner on first base

Sample space:
all possible outcomes of an experiment.
Denoted by S.

Event:
any subset of the sample space S.
Typically denoted A, B, C, etc.
Null event: the empty set \( \Phi \)
Certain event: S

EXAMPLES
1. Toss a coin one time. \( S = \{ H, T \}; A = \{ H \}, B = \{ T \} \)
2. Toss a die one time and count the dots on the upper face. \( S = \{ 1, 2, 3, 4, 5, 6 \} \)
   \( A = \) even number of dots on upper face = \{ 2, 4, 6 \}
   \( B = 3 \) or fewer dots on upper face = \{ 1, 2, 3 \}
3. Attempt a field goal.
   \( S = \{ \) kick is successful, kick fails, kick is blocked, holder fumbles snap \}
LAWS OF PROBABILITY

1) \[ 0 \leq P(A) \leq 1 \] for any event \( A \)

2) \[ P(\emptyset) = 0 \text{ and } P(S) = 1 \]

For any event \( A \), \( A' \) is the complement of \( A \); \( A' \) is everything in the sample space \( S \) that is not in \( A \).

3) \[ P(A') = 1 - P(A) \] for any event \( A \)

EXAMPLE:

Toss a coin one time. \( S = \{H, T\} \); \( P(H) = 1 - P(T) \).

Birthday Problem.

Two events \( A \) and \( B \) are disjoint (or mutually exclusive) if they have no outcomes in common and so can never simultaneously.

Addition Rule for Disjoint Events:

4) If \( A \) and \( B \) are disjoint events,

\[ P(A \cup B) = P(A) + P(B) \]

Note:
1. \( \cup \) means “union”; \( \cap \) means “intersection”.
2. Read “\( A \cup B \)” as “\( A \) or \( B \) or both”; read “\( A \cap B \)” as “\( A \) and \( B \)”.
3. Events \( A \) and \( B \) are disjoint (or mutually exclusive) if they do not intersect, that is, if \( A \cap B = \emptyset \)

Venn diagram

General Addition Rule

5) For any two events \( A \) and \( B \), the probability that one or the other occurs is

\[ P(A \text{ or } B) = P(A \cup B) = P(A) + P(B) - P(A \cap B) \]

Venn diagram

Multiplication Rule

6) For two independent events \( A \) and \( B \), the probability that \( A \) and \( B \) occur is

\[ P(A \text{ and } B) = P(A \cap B) = P(A) \times P(B) \]
EXAMPLE: At a particular college 56% of all students live on campus, 62% of all students purchased a campus meal plan, and 42% do both.

QUESTION: what is the probability that a randomly selected student either lives or eats on campus?
Let \( L = \{ \text{student lives on campus} \} \), \( M = \{ \text{student purchased a campus meal plan} \} \).

\[
P(\text{a student either lives or eats on campus}) = P(L \text{ or } M) = P(L) + P(M) - P(L \text{ and } M) = .56 + .62 - .42 = .76
\]

EXAMPLE: (To bunt or not to bunt)

The table below lists the probabilities of scoring at least one run in situations that are defined by the number of outs and the bases occupied. These probabilities are determined from the analysis of thousands of games and game situations in the American League. For example, the probability of scoring at least one run when there are no outs and a runner on first base .39.

<table>
<thead>
<tr>
<th>BASES OCCUPIED</th>
<th>Bases empty</th>
<th>1st base</th>
<th>2nd base</th>
<th>3rd base</th>
<th>1st, 2nd</th>
<th>1st, 3rd</th>
<th>2nd, 3rd</th>
<th>1st, 2nd, 3rd</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 out</td>
<td>.26</td>
<td>.39</td>
<td>.57</td>
<td>.72</td>
<td>.59</td>
<td>.76</td>
<td>.83</td>
<td>.81</td>
</tr>
<tr>
<td>1 out</td>
<td>.16</td>
<td>.26</td>
<td>.42</td>
<td>.55</td>
<td>.45</td>
<td>.61</td>
<td>.74</td>
<td>.67</td>
</tr>
<tr>
<td>2 outs</td>
<td>.07</td>
<td>.13</td>
<td>.24</td>
<td>.28</td>
<td>.24</td>
<td>.37</td>
<td>.37</td>
<td>.43</td>
</tr>
</tbody>
</table>

Let’s concentrate on the strategy of the sacrifice bunt. The purpose of the sacrifice bunt is to “sacrifice” the batter and make an “out” to advance the baserunner(s) to the next base. It can be utilized when there are less than 2 outs and a baserunner.

Ignoring the suicide squeeze bunt and other low probability outcomes, the following four outcomes can occur as the result of a sacrifice bunt:

**Outcome 1:** the bunt is successful; the runner or runners advance one base and the batter is out.

**Outcome 2:** the bunt is not successful; the bunt results in an “out” and runner(s) is(are) not advanced.

**Outcome 3:** the batter bunts into a double play

**Outcome 4:** the batter reaches first base safely and the runner advances one base.

You are the manager of an American League team playing a game that is tied in the 7th inning. Your team has a runner on first base with no one out. Should you order your batter to attempt a sacrifice bunt if the probabilities of the 4 outcomes are as shown below.

\[
P(\text{outcome 1}) = .75; \ P(\text{outcome 2}) = .10, \ P(\text{outcome 3}) = .10, \ P(\text{outcome 4}) = .05
\]

Current probability of scoring at least one run:
Probability of scoring at least one run if sacrifice bunt:
EXAMPLE: (no-hitters - normal model).
In baseball a no-hitter is a regulation nine-inning game in which the pitcher does not allow the opposing team to get any hits. Based on a statistical analysis of MLB games, the number of hits yielded per team per game can be approximated very well with a normal model with mean 8.72 and standard deviation 1.10.

What is the probability of a no-hitter? z-score for 0 hits =

EXAMPLE: (no-hitters again - relative frequency approach).
Since 1900 there have been approximately 183,000 regular season games and 243 no-hitters. We can use the relative frequency approach to approximate the probability of a no-hitter. Since each game represents 2 opportunities for a no-hitter,

\[ P(\text{no-hitter}) = \frac{243}{183,000} \]

EXAMPLE: (independent events)
In 1938 Johnny Vander Meer pitched 2 no-hitters in a row. If we use the probability of a no-hitter from the previous example and assume games are independent, then

\[ P(2 \text{ no-hitters in a row}) = \left( \frac{243}{183,000} \right)^2 \]

EXAMPLE: (Great Collapses in Sports)
In the 2005 NCAA basketball tournament (eventually won by UNC when they beat Illinois), Arizona led Illinois 75-60 with 4:05 left in the championship game of the Chicago regional. Illinois won the game. Can we approximate the probability of this collapse by Arizona?

When discussing the normal model we observed that the difference in points for the 2 teams in an NCAA game follows a normal model with standard deviation 10 points. For 2 evenly matched teams we can assume mean difference = 0.

The standard deviation for a 4:05 time segment is 3.2 points (later we'll learn how to calculate this).

\[ P(\text{Arizona scores 15 points less than Illinois in last 4:05}) = \text{NORMDIST}(-15, 0, 3.2, \text{TRUE}) = 0.00000138 \] or about 1 in 724,638.
Tree Diagrams

**EXAMPLE (Probability of playing professional baseball)**

It is known that 6.1% of high school baseball players go on to play baseball at the college level. Of these, 9.4% will play professionally. Unlike football and basketball, high school players can also go directly to professional baseball without playing in college. Studies have shown that given that a high school player does not compete in college, the probability he plays professionally is .002.

**Question 1:** What is the probability that a high school baseball player ultimately plays professional baseball?

![Tree Diagram for Playing Professional Baseball]

**Question 2:** Given that a high school baseball player played professionally, what is the probability he played in college?

---

**Tree Diagram for Rare Disease Testing**

**EXAMPLE : (Diagnostic Test for AIDS)**

\[ V = \{ \text{person has HIV} \}; \quad \text{based on CDC estimates,} \quad P(V) = \frac{1.5\text{ million}}{250\text{ million}} = .006 \]

Define the following events:
- \(+ : \) positive test (test indicates HIV present)
- \(- : \) test outcome is negative

The following reliabilities for the particular HIV test have been determined by clinical trials:
- i) When HIV is present, the test is positive 99.9% of the time
- ii) When the disease is absent, the test is negative 99% of the time

**Question 1:**
- What is the probability that a randomly selected person will test positive?
Answer: There are 2 sequences of branches that lead to a positive test. To find the probability of a positive test, add the probabilities of these 2 sequences of branches:

\[ P(\text{test positive}) = \]

Question 2: Given that a person has tested positive, what is the probability that the person actually has HIV?

Answer: The 2 sequences of branches that lead to a positive test have probabilities that add to .01593. Only 1 of these sequences represented people that actually have HIV (this sequence has probability .00599). So our probability of interest is

\[ P(\text{person has HIV given that tested positive}) = \]

HOW CAN THIS BE?

---

### Relationship Between Odds and Probabilities

\[ \Rightarrow \text{From Probabilities to Odds:} \]

If event A has probability \( P(A) \), then the odds in favor of A are \( P(A) \) to 1-\( P(A) \). It follows that the odds against A are 1-\( P(A) \) to \( P(A) \)

**EXAMPLE:**

If the probability of an earthquake in California is .25, then the odds in favor of an earthquake are .25 to .75 or 1 to 3. The odds against an earthquake are .75 to .25 or 3 to 1.

\[ \Rightarrow \text{From Odds to Probabilities:} \]

If the odds in favor of an event E are \( a \) to \( b \), then

\[ P(E) = \frac{a}{a + b} \]

in addition,

\[ P(E') = \frac{b}{a + b} \]

**EXAMPLE:**

If the odds in favor of Duke winning the NCAA’s are 3 (a) to 1 (b), then

\[ P(\text{Duke wins}) = \frac{3}{3 + 1} = \frac{3}{4} \]
in addition,

\[ P(\text{Duke does not win}) = \frac{1}{3 + 1} = \frac{1}{4} \]

---

### Probability Models

One way of assigning probabilities: **Equally Likely** approach

- if an experiment has \( n \) outcomes, then each outcome has probability \( \frac{1}{n} \) of occurring

if an event \( A_1 \) has \( n_1 \) outcomes, \( P(A_1) = \frac{n_1}{n} \)

**Factorial Notation**

\[ n! = \]

**Examples**

**Permutations**

**Combinations**

To calculate factorials, permutations and combinations using **Excel**, use the “paste function” by clicking on the icon \( f_x \).

To calculate factorials, permutations and combinations using the **TI calculator**: **Calculator Appendix, p. 10.**

**Examples**: (state lotteries; NBA draft lottery)

---

**News Clip**

Heads I win, tails you lose.

— 17th century English saying