Homework (due Wed, Oct 27)
Chapter 7: #17, 27, 28

Announcements:

- Midterm exams keys on web. (For a few hours the answer to MC#1 was incorrect on Version A.)
- *No* grade disputes now. Will have a chance to do that in writing at end of quarter.
- Grades are curved at end of quarter, not now.
- Material gets harder from here on! Practice problems with answers will be posted on website.

Chapter 7
Probability

Today: 7.1 to 7.3, small part of 7.4

Fri: Finish 7.4, 7.5

Skip Section 7.6

Mon: Section 7.7 and supplemental material on intuition and probability

Random Circumstance

A *random circumstance* is one in which the *outcome* is *unpredictable*.

Could be unpredictable because:

- It *isn't determined* yet

  or

- We have *incomplete knowledge*

Example of a random circumstance

- Sex of an unborn child is unpredictable, so it is a random circumstance.
- We can talk about the *probability* that a child will be a boy.

Why is it unpredictable?

- Before conception:
  - It *isn't determined* yet
- After conception:
  - We have *incomplete knowledge*

Goals in this chapter

- Understand what is meant by “probability”
- Assign probabilities to possible outcomes of random circumstances.
- Learn how to use probability wisely

What does probability mean??

What does it mean to say:

- *The probability of rain tomorrow is .3.*
- *The probability that a coin toss will land heads up is ½.*
- *The probability that humans will survive to the year 3000 is .8.*

Is the word "probability" interpreted the same way in all of these?

Two basic interpretations of probability (Summary box, p. 237)

Interpretation 1: Relative frequency

- Used for *repeatable* circumstances
- The *probability* of an outcome is the *proportion* of time that outcome does or will happen *in the long run.*

Interpretation 2: Personal probability (subjective)

- Most useful for *one-time events*
- The *probability* of an outcome is *the degree to which* an individual *believes* it will happen.

Two methods for determining relative frequency probability

1. **Make an assumption** about the physical world *or*
2. **Observe** the *relative frequency* of an outcome over many repetitions. "Repetitions" can be:
   a. Over *time*, such as how often a flight is late
   b. Over *individuals*, by measuring a representative sample from a larger population and observing the *relative frequency* of an outcome or category of interest, such as the probability that a randomly selected person has a certain gene.

How relative frequency probabilities are determined, Method 1:

Make an assumption about the physical world. Examples:

- Flip a coin, **probability** it lands heads = ½.

We *assume* the coin is balanced in such a way that it is equally likely to land on either side.

- Draw a card from a shuffled, regular deck of cards, **probability** of getting a heart = ¼

We *assume* all cards are equally likely to be drawn

How relative frequency probabilities are determined, Method 2a (over time):

Observe the *relative frequency* of an outcome over many repetitions (*long run relative frequency)*

Probability that a flight will be on time:

- According to United Airline's website, the probability that Flight 436 from SNA to Chicago will be *on time* (within 14 minutes of the stated time) is 0.90.
- Based on *observing* this particular flight over many, many days; it was on time on 90% of those days. The *relative frequency* on time = .9

How relative frequency probabilities are determined, Method 2b (over individuals):

Measure a representative sample and observe the *relative frequency* of possible outcomes or categories for the sample

Probability that an adult female in the US believes in life after death is about .789. [It's .72 for males]

- Based on a national survey that asked 517 women if they believed in life after death
- 408 said yes
- Relative frequency is 408/517 = .789

Note about methods 2a and 2b:

Usually these are just *estimates* of the true probability, based on *n* repetitions or *n* people in the sample. So, they have an associated *margin of error* with them.

Example:

Probability that an adult female in the US believes in life after death = .789, based on n = 517 women.

Margin of error is $\frac{1}{\sqrt{517}}$ = .044.

Personal Probability
Especially useful for one-time only events

The *personal probability* of an outcome is *the degree to which* an individual *believes* it will happen.

Examples:

- What is the probability that *you* will get a B in this class? We can't base the answer on relative frequency!
- LA Times, 10/8/09, scientists have determined that the probability of the asteroid Apophis hitting the earth in 2036 is 1 in 250,000. In 2004, they thought the probability was .027 that it would hit earth in 2029. "Expert opinion" has been updated.

Notes about personal probability

- Sometimes the individual is an expert, and combines subjective information with data and models, such as in assessing the probability of a magnitude 7 or higher earthquake in our area in the next 10 years.
- Sometime there is overlap in these methods, such as determining probability of rain tomorrow – uses similar pasts.

Clicker questions *not* for credit!

The probability that the winning "Daily 3" lottery number tomorrow evening will be 777 is 1/1000.

Which interpretation is best?

A. ***Rel. freq. based on physical assumption***

B. Rel. freq. based on long run over time

C. Rel. freq. based on representative sample

D. Personal probability

Clicker questions *not* for credit!

The probability that the SF Giants will win the World Series this year is .40.

Which interpretation is best?

A. Rel. freq. based on physical assumption

B. Rel. freq. based on long run over time

C. Rel. freq. based on representative sample

D. Personal probability

Clicker questions *not* for credit!

The probability that the plaintiff in a medical malpractice suit will win is .29. (Based on an article in *USA Weekend)*

Which interpretation is best?

A. Rel. freq. based on physical assumption

B. Rel. freq. based on long run over time*

C. Rel. freq. based on representative sample*

D. Personal probability

* We would need more information to know which of these is correct.

Section 7.3: Probability definitions and relationships

We will use 2 examples to illustrate:

1. Daily 3 lottery winning number.

   Outcome = 3 digit number, from 000 to 999

2. Choice of 3 parking lots, you always try lot 1, then lot 2, then lot 3.

   Lot 1 works 30% of the time, you aren't late

   Lot 2 works 50% of the time, you are late

   Lot 3 always works, so you park there 20% of the time, and when you do, you are very late!

Definitions of *Sample space* and *Simple event*

The sample space S for a random circumstance is the collection of unique, non-overlapping outcomes.

A simple event is *one outcome* in the sample space.

Ex 1: S = {000, 001, 002, …, 999}

Simple event: 659

There are 1000 simple events.

Ex 2: S = {Lot 1, Lot 2, Lot 3}

Simple event: Lot 2

There are 3 simple events.

Definition and notation for an *event*

Definition: An event is *any subset* of the sample space. (One or more simple events)

Notation: A, B, C, etc.

Ex 1: A = *winning number begins with 00*

A = {000, 001, 002, 003,…, 009}

B = *all same digits* = {000, 111, …, 999}

Ex 2: A = *late for class* = {Lot 2, Lot 3}

Probability of Events: Notation and Rules (for all interpretations and methods)

Notation: P(A) = probability of the event A

Rules: Probabilities are always assigned to *simple events* such that these 2 rules must hold:

1. $0 \leq P(A) \leq 1$ for each simple event A
2. The *sum* of probabilities of *all* simple events in the sample space is 1.

The probability of *any* event is the sum of probabilities for the simple events that are part of it.

Special Case: Assigning Probabilities to *Equally Likely* Simple Events

P(A) = probability of the event A

Remember, Conditions for Valid Probabilities:

- Each probability is between 0 and 1.
- The sum of the probabilities over all possible simple events is 1.

Equally Likely Simple Events

If there are k simple events in the sample space and they are all equally likely, then the probability of the occurrence of each one is $1/k$.

Example: California Daily 3 Lottery

Random Circumstance:

A three-digit winning lottery number is selected.

Sample Space: {000, 001, 002, 003, . . . , 997, 998, 999}.

There are 1000 simple events.

Probabilities for Simple Event: Probability that any specific three-digit number is a winner is 1/1000.

Physical assumption: all three-digit numbers are equally likely.

Event A = last digit is a 9 = {009, 019, . . . , 999}.

***P*(A) = 100/1000 = 1/10**.

Event B = three digits are all the same

= {000, 111, 222, 333, 444, 555, 666, 777, 888, 999}.

Since event B contains 10 events, ***P*(B) = 10/1000 = 1/100**.

Example 2:
Simple events are *not* equally likely

Simple Event	*Probability*
Park in Lot 1	.30
Park in Lot 2	.50
Park in Lot 3	.20

Note that these sum to 1

Event A = late for class = {Lot 2, Lot 3}

P(A) = .50 + .20 = .70

Probability in daily language:

People often express probabilities as percents, proportions, probabilities:

- United flight 436 from SNA to Chicago is late 10 percent of the time.
- The proportion of time United flight 436 is late is .1.
- The probability that United flight 436 from SNA to Chicago will be late is .1.

These are all equivalent.

RELATIONSHIPS BETWEEN EVENTS

- Defined for events in the *same random circumstance only:*
  - Complement of an event
  - Mutually exclusive events = disjoint events
- Defined for events in the same *or* different random circumstances:
  - Independent events
  - Conditional events

Definition and Rule 1 (apply to events in the *same* random circumstance):

Definition: One event is the **complement** of another event if:

- They have no simple events in common, AND
- They cover all simple events

Notation: The complement of A is $\mathbf{A^C}$

$$\boxed{\textbf{RULE 1: } P(A^C) = 1 - P(A)}$$

Ex 2: *Random circumstance* = parking on one day

A = late for class, A^C = on time

$P(A) = .70$, so $P(A^C) = 1 - .70 = .30$

Complementary Events, Continued

Rule 1: $P(\mathrm{A}) + P(\mathrm{A}^C) = 1$

Example: *Daily 3 Lottery*

A = player buying single ticket wins

A^C = player does not win

$P(\mathrm{A}) = 1/1000$ so **$P(\mathrm{A}^C) = 999/1000$**

Example: *On-time flights*

A = flight you are taking will be on time

A^C = flight will be late

Suppose $P(\mathrm{A}) = .80$, then **$P(\mathrm{A}^C) = 1 - .80 = .20$.**

Mutually Exclusive Events

Two events are **mutually exclusive**,
or equivalently **disjoint**, if they do not contain *any* of the same simple events (outcomes).
(Applies in *same random circumstance*.)

Example: Daily 3 Lottery

A = all three digits are the same (000, 111, etc.)

B = the number starts with 13 (130, 131, etc.)

The events A and B are mutually exclusive (disjoint), but they are not complementary.
(No overlap, but *don't* cover all possibilities.)

Independent and Dependent Events

- Two events are **independent** of each other if knowing that one will occur (or has occurred) *does not change* the probability that the other occurs.
- Two events are **dependent** if knowing that one will occur (or has occurred) *changes* the probability that the other occurs.

The definitions can apply *either* ...
to events *within the same random circumstance* or
to events *from two separate random circumstances.*

EXAMPLE OF INDEPENDENT EVENTS

- Events in the *same random circumstance:*

  Daily 3 lottery on the *same* draw

  A = *first* digit is 0

  B = *last* digit is 9

  Knowing first digit is 0, P(B) is *still* 1/10.
- Events in *different random circumstances*:

  Daily 3 lottery on *different* draws

  A = *today's* winning number is 191

  B = *tomorrow's* winning number is 875

  Knowing today's # was 191, P(B) is *still* 1/1000

Mutually exclusive or independent?

- If two events are mutually exclusive (disjoint), they *cannot* be independent:
  - If *disjoint*, then knowing A occurs means P(B) = 0
  - In *independent*, knowing A occurs gives no knowledge of P(B)

Example of mutually exclusive (disjoint):

A = *today's* winning number is 191,

B = *today's* winning number is 875

Example of independent:

A = *today's* winning number is 191

B = *tomorrow's* winning number is 875

Conditional Probabilities

The conditional probability of the event B, given that the event A will occur or has occurred, is the long-run relative frequency with which event B occurs when circumstances are such that A also occurs; written as $\boldsymbol{P(B|A)}$.

$P(B)$ = *unconditional* probability event B occurs.

$P(B|A)$ = "probability of B given A"

= *conditional* probability event B occurs *given that we know A has occurred or will occur*.

EXAMPLE OF CONDITIONAL PROBABILITY

TABLE 2.3 Nighttime Lighting in Infancy and Eyesight

Slept with:	No Myopia	Myopia	High Myopia	Total
Darkness	$\boxed{155}$ (90%)	15 (9%)	2 (1%)	$\boxed{172}$
Nightlight	153 (66%)	72 (31%)	7 (3%)	232
Full Light	34 (45%)	36 (48%)	5 (7%)	75
Total	$\boxed{342}$ (71%)	123 (26%)	14 (3%)	$\boxed{479}$

Random circumstance: Observe one randomly selected child

A = child slept in darkness as infant [Use "total" column.]

P(A) = 172/479 = .36

B = child did not develop myopia [Use "total" row]

P(B) = 342/479 = **.71**

P(B|A) = P(no myopia | slept in dark) [Use "darkness" row]

= 155/172 = **.90** ≠ **P(B)**

NOTES ABOUT CONDITIONAL PROBABILITY

1. P(B|A) generally does *not* equal P(B).
2. P(B|A) = P(B) *only* when A and B are independent events
3. In Chapter 6, we were actually testing if two types of events were *independent.*
4. *Conditional* probabilities are similar to *row* and *column* proportions (percents) in contingency tables. (Myopia example on previous page.)