Announcements:

- Slight change in schedule: We will cover 8.1 to 8.3 on Monday, then 7.7 on Wed, because you need 8.1 to 8.3 for Discussion.
- If you haven't taken the survey yet, I encourage you to do it this weekend!
- Chapter 7 practice problems are now on website.

Homework (due Wed, Oct 27th):

Chapter 7:

#35, 51, 53 (answer in back, show work), 55

Chapter 7

Probability continued: Finish slides from Wed first

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 ***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	155 (90%)	15 (9%)	2 (1%)	172
Nightlight	153 (66%)	72 (31%)	7 (3%)	232
Full Light	34 (45%)	36 (48%)	5 (7%)	75
Total	342 (71%)	123 (26%)	14 (3%)	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.)

Review:
RELATIONSHIPS BETWEEN EVENTS

- Defined for events in the *same random circumstance only:*
  - **Complement** of an event **A** is all simple events *not* part of **A.**
  - **Mutually exclusive = disjoint events** have no overlapping simple events
  - **Complements are disjoint events** (no overlapping simple events)
  - **Disjoint events aren't complements** *unless* they cover all possibilities.

Review: 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.*

HOW TO DETERMINE IF TWO EVENTS ARE INDEPENDENT

1. Physical assumption
   Example: Lottery draws on different days don't affect each other.
2. See if P(B|A) = P(B). If so, A and B are independent events, otherwise they are not.
   Example: Suppose data showed that smokers and non-smokers are equally likely to get the flu.
   P(flu | smoker) = P(flu) so they are independent.
3. Events A and B are independent if and only if P(A and B) = P(A)P(B). Sometimes have info so you can check to see if this is true.

Testing independence of myopia and lighting

TABLE 2.3 Nighttime Lighting in Infancy and Eyesight

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

A = child slept in darkness as infant
P(A) = 172/479 = .36
B = child did not develop myopia
P(B) = 342/479 = .71

Does P(A and B) = P(A)P(B) = (.36)(.71) = .2556?

P(A and B) = P(darkness and no myopia) = 155/479 = .3236

NO, P(A and B) ≠ P(A)P(B), so A and B are *not* independent.

Infant sleeptime lighting and myopia are not independent.

7.4 Basic Rules for Finding Probabilities

Probability an Event Does Not Occur

Rule 1 (for "not the event"): $P(A^C) = 1 - P(A)$

Example: In the US, blood type percents* are:
O: 46%, **A**: 39%, **B**: 11%, **AB**: 4%
Suppose we randomly select someone.
P(Blood type O) = 0.46
so *P*(Blood type A, B, or AB) = 1 – 0.46 = 0.54.

*Source: www.bloodbook.com

Probability That Either One or Both of Two Events Happen

Rule 2 (addition rule for "either/or/both"):

Rule 2a (general):
P(A *or* B) = *P*(A) + *P*(B) – *P*(A and B)

Rule 2b (for mutually exclusive events):
If A and B are *mutually exclusive* events,
P(A or B) = *P*(A) + *P*(B)

Illustrate with Venn diagram on board.

Example 7.14 *Roommate Compatibility*

Brett is off to college and will be assigned a roommate from 1000 other males. Will he get one who snores? Parties? Both?

		Snores? Yes	No	Total
Likes to Party?	Yes	150	100	250
	No	200	550	750
		350	650	1000

A = likes to party P(A) = **250**/1000 = 0.25

P(A and B) = **150**/1000 = .15

B = snores P(B) = **350**/1000 = 0.35

Probability Brett will be assigned a roommate who either likes to party *or* snores, *or* both is: ***P*(A *or* B)**
= ***P*(A)** + ***P*(B)** – ***P*(A and B)** = 0.25 + 0.35 – 0.15 = 0.45
So the probability his roommate is acceptable is 1 – 0.45 = 0.55

Probability That Two or More Events Occur Together

Rule 3 (multiplication rule for "and"):

Rule 3a (general):

$P(\text{A } and \text{ B}) = P(\text{A})P(\text{B}|\text{A})$

Rule 3b (for independent events):

If A and B are *independent* events,

$P(\text{A } and \text{ B}) = P(\text{A})P(\text{B})$

Extension of Rule 3b (> 2 independent events):

For several independent events,

$P(\text{A}_1 \text{ } and \text{ A}_2 \text{ } and \ldots and \text{ A}_n) = P(\text{A}_1)P(\text{A}_2)\ldots P(\text{A}_n)$

Example with *independent* events (Rule 3b)

The probability of a birth being a boy is .512. Suppose a woman has 2 kids.

A = first child is a boy

B = second child is a boy

We assume these are independent events.

P(**A** *and* **B**) = P(**A**)P(**B**) = (.512)(.512) = .2621

Probability of 2 girls is (.488)(.488) = .2381

From **Rule 2b**, probability of same sex = .5002

From **Rule 1**, probability of different sex = .4998

Example with *dependent* events (Rule 3a)

Randomly select a student who takes Stat 8 with me.

A = student comes to class regularly; P(**A**) = 0.7

A^C = student doesn't come regularly; P(**A**^{C}) = 0.3

B = student gets at least a B in the course

P(**B**|**A**) = 0.8   P(**B**|**A**^{C}) = 0.4

What is the probability that the student comes to class regularly *and* gets at least a B?

P(**A** and **B**) = *P*(**A**)*P*(**B**|**A**) = (.7)(.8) = .56

Extension of Rule 3b: *Blood Types Again*

A blood bank needs Type O blood. What is the probability that the next 3 donors (not related) all have Type O blood? We can assume independent.

Event A = 1st donor has Type O blood, *P*(A) = .46
Event B = 2nd donor has Type O blood, *P*(B) = .46
Event C = 3rd donor has Type O blood, *P*(C) = .46

P(Next 3 donors all have Type O blood)
= P(A *and* B *and* C)
= (.46)(.46)(.46) = $(.46)^3$ = .097336

Determining a Conditional Probability

Rule 4 (conditional probability):

$$P(\text{B}|\text{A}) = \frac{P(A \text{ and } B)}{P(A)}$$

$$P(\text{A}|\text{B}) = \frac{P(A \text{ and } B)}{P(B)}$$

EXAMPLE OF CONDITIONAL PROBABILITY - Revisited

TABLE 2.3 Nighttime Lighting in Infancy and Eyesight

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

A = child slept in darkness as infant

P(A) = 172/479 = .36

B = child did not develop myopia

P(B) = 342/479 = .71

P(B|A) = P(no myopia | slept in dark) = 155/172 = .90

But this is $\frac{P(A \text{ and } B)}{P(A)} = \frac{155/479}{172/479} = \frac{155}{172}$

Mutually exclusive versus Independent

Students sometimes confuse the definitions of **independent** and **mutually exclusive** events.

- When two events are *mutually exclusive* and one happens, it *turns the probability of the other one to **0***.
- When two events are *independent* and one happens, it *leaves the probability of the other one alone*.

In Summary (see page 249) …

When Events Are:	P(A *or* B) is:	P(A *and* B) Is:	P(A|B) Is:
Mutually Exclusive	$P(A) + P(B)$	0	0
Independent	$P(A) + P(B) - P(A)P(B)$	$P(A)P(B)$	$P(A)$
Any	$P(A) + P(B) - P(A \text{ and } B)$	$P(A)P(B \mid A)$	$\frac{P(A \text{ and } B)}{P(B)}$

7.5 Strategies for Finding Complicated Probabilities

Example 7.20 *Winning the Daily 3 Lottery*
Event A = winning number is 956. **What is *P*(A)?**

Method 1: With physical assumption that all 1000 possibilities are equally likely, $P(A) = 1/1000$.

Method 2: Define three events,
B_1 = 1st digit is 9, B_2 = 2nd digit is 5, B_3 = 3rd digit is 6
Event A occurs if and only if all 3 of these events occur.
Note: $P(B_1) = P(B_2) = P(B_3) = 1/10$. Since these events are all *independent*, we have $P(A) = (1/10)^3 = 1/1000$.

**** Can be more than one way to find a probability.***

Steps for Finding Probabilities

Step 1: List each separate random circumstance involved in the problem.

Step 2: List the possible outcomes for each random circumstance.

Step 3: Assign whatever probabilities you can with the knowledge you have.

Step 4: Specify the event for which you want to determine the probability.

Step 5: Determine which of the probabilities from Step 3 and which probability rules can be combined to find the probability of interest.

Some Hints for Finding Probabilities

- ***P*(A and B):** Sometimes you can define the event in physical terms and know the probability or find it from a two-way table.

  Example: I could classify the class into male, female and also year in school. Then, for example, probability that a randomly selected student in the class is Male *and* sophomore is the proportion of the class in that cell of the table. Don't need separate P(A) and P(B).

- Check if probability of the ***complement*** is **easier** to find, then subtract it from 1 (applying Rule 1).

  Example: Probability of at least 1 boy in family of 3 kids =
  1 – Probability of all girls = $1 - (.488)^3 = 1 - .116 = .884$

Finding Conditional Probability in Opposite Direction

Know *P*(B|A) but want *P*(A|B): Use Rule 3a to find $P(B) = P(A \text{ and } B) + P(A^C \text{ and } B)$, then use Rule 4.

$$P(A \mid B) = \frac{P(A \text{ and } B)}{P(B \mid A)P(A) + P(B \mid A^C)P(A^C)}$$

Let's look at some tools that are much easier than using this formula!

Solve these probability questions

Example 1: Medical testing for a rare disease

D = person has the disease, suppose:

$P(D) = 1/1000 = .001$, $P(D^C) = .999$

T = test for the disease is positive, suppose:

$P(T \mid D) = .95$, so $P(T^C \mid D) = .05$

$P(T \mid D^C) = .05$, so $P(T^C \mid D^C) = .95$

So the test is ***95% accurate*** whether person has the disease or not

Find **P(D | T)**

= Probability of disease, *given* the test is positive

Example 2: Probability of getting a B or better

Return to example of Statistics 8 grades

A = student comes to class regularly; **P(A)** = 0.7

B = student gets at least a B in the course

P(B|A) = 0.8 **P(B|A^C)** = 0.4

Question:

What is the overall probability of getting at least a B?

Two-Way Table: "Hypothetical Hundred Thousand"

Table of hypothetical 100,000 people who get tested

1/1000 of them have disease = 100 people

Of those, 95% = **95** people test positive, so 5 negative

999/1000 of them don't have the disease = 99,900 people

Of those, 95% = 94,905 people test *negative*, so 4995 positive

Read from Table: P(Disease | Test positive) = 95/5090 = .019

	Test pos	Test neg	
Disease	**95**	5	100
No disease	4995	94,905	99,900
	5090	94,910	100,000

Some definitions from Section 7.7 Probability of accurate medical test

Define the events:

D = person has the disease

D^C = person does not have the disease

T = test is positive

T^C = test is negative

Sensitivity of a test = $P(T \mid D)$, i.e., correct outcome if person *has* the disease.

Specificity of a test = $P(T^C \mid D^C)$, i.e. correct outcome if person *does not have* the disease.

Tree Diagrams

Step 1: Determine first random circumstance in sequence, and create first set of branches for possible outcomes. Create one branch for each outcome, write probability on branch.

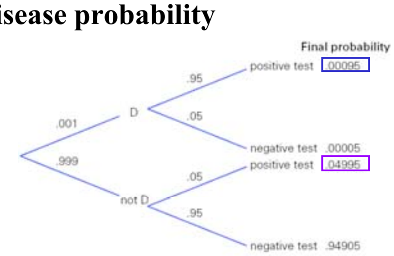
Step 2: Determine next random circumstance and append branches for possible outcomes to each branch in step 1. Write associated *conditional probabilities* on branches.

Step 3: Continue this process for as many steps as necessary.

Step 4: To determine the probability of following any particular sequence of branches, multiply the probabilities on those branches. This is an application of Rule 3a.

Step 5: To determine the probability of any collection of sequences of branches, add the individual probabilities for those sequences, as found in step 4. This is an application of Rule 2b.

Disease probability



Sensitivity = specificity = 0.95

P(D *and* positive test) = (.001)(.95) = .00095.

P(test is positive) = .00095 + .04995 = .0509.

P(D | positive test) = .00095/.0509 = .019

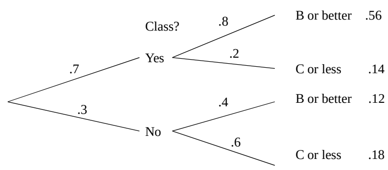
Example 2: Return to example of Stat 8 grades

A = student comes to class regularly; $P(A) = 0.7$, $P(A^C) = 0.3$

B = student gets at least a B in the course

$\mathbf{P(B|A)} = 0.8$ $\mathbf{P(B|A^C)} = 0.4$

Question: What is the overall probability of getting at least a B?



So, probability of B or better = .56 + .12 = .68 overall

More examples on board, if time:

1. Suppose there is no relationship between two variables, e.g. listening to Mozart and increased IQ. Suppose 3 independent experiments are done, each using the 0.05 criterion for statistical significance.

What is the probability that *at least one* finds statistical significance just by chance?

2. Suppose you drive on a certain freeway daily. Speed limit is 65. You drive over 65 all the time, but over 75 about 30% of the time.

P(ticket | over 75) = 1/50 = .02

P(ticket | 65 to 75) = 1/200 = .005.

What is the probability you get a ticket on a randomly selected day?

Problem 1

Suppose there is no relationship between two variables, e.g. listening to Mozart and increased IQ. Suppose 3 independent experiments are done, each using the 0.05 criterion for statistical significance.

What is the probability that *at least one* experiment results in a statistically significant relationship just by chance?

Problem 2

Suppose you drive on a certain freeway daily. Speed limit is 65. You drive over 65 all the time, but over 75 about 30% of the time.

P(ticket | over 75) = 1/50 = .02

P(ticket | 65 to 75) = 1/200 = .005.

What is the probability you get a ticket on a randomly selected day?