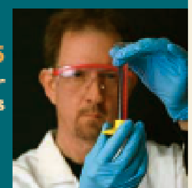
PROBABILITY IN OUR DAILY LIVES

Chapter 5

PART 2

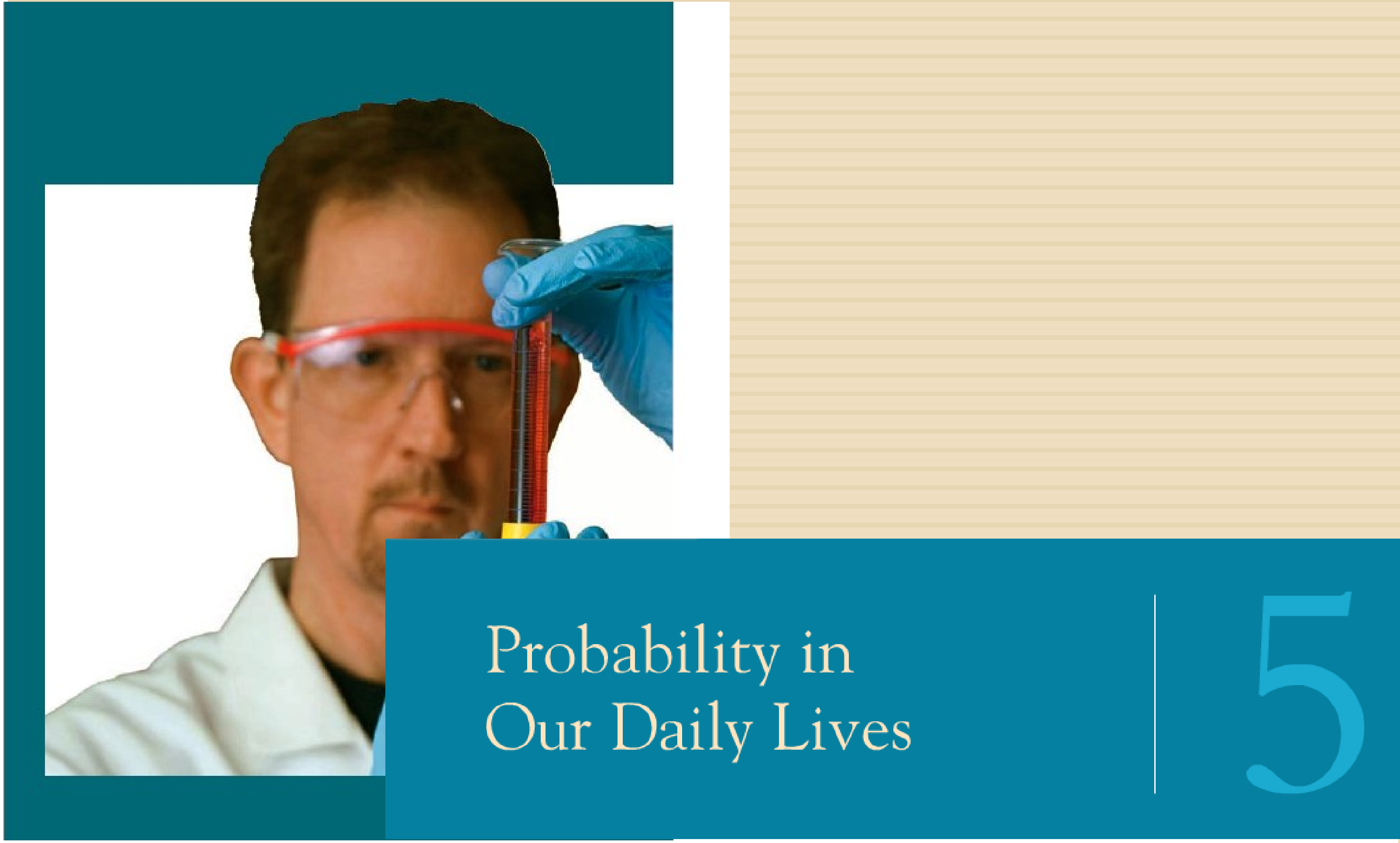
Probability, Probability Distributions and Sampling Distributions

Chapter 5
Probability in Our Daily Lives



Chapter 6
Probability Distributions

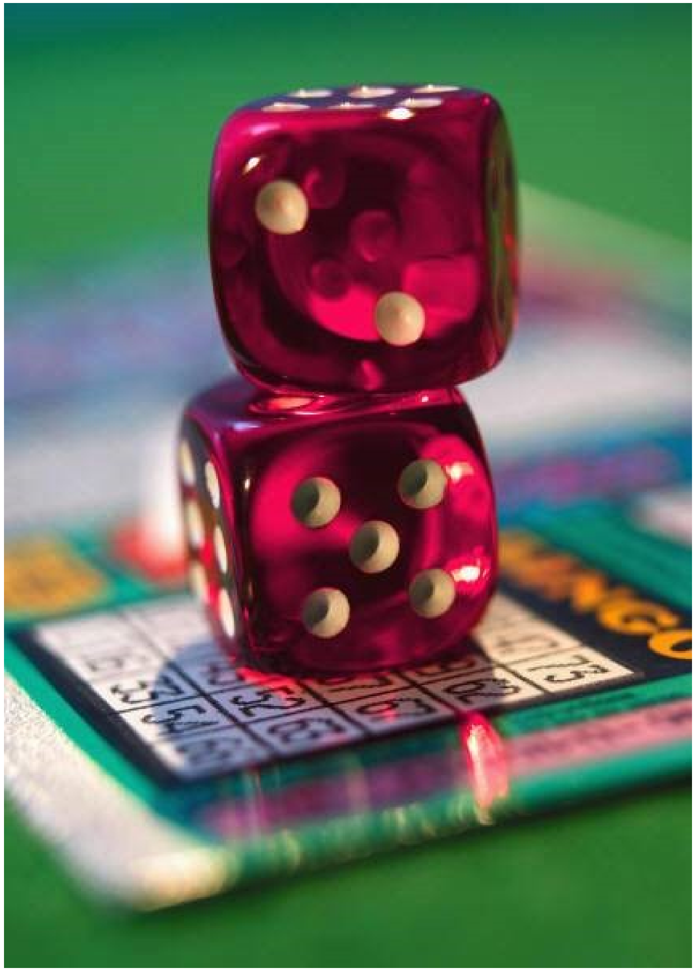
Chapter 7
Sampling Distributions



5.1 How Can Probability Quantify Randomness?

Learning Objectives

1. Random Phenomena
2. Law of Large Numbers
3. Probability
4. Independent Trials
5. Finding probabilities
6. Types of Probabilities: Relative Frequency and Subjective



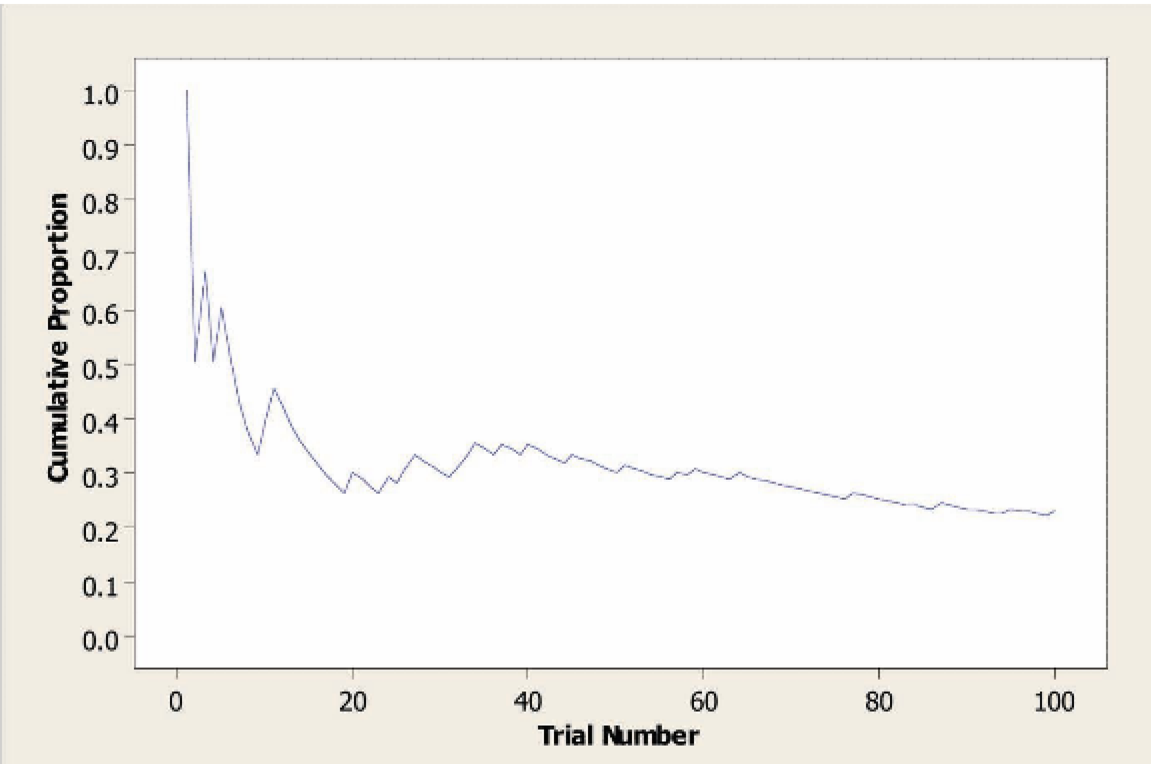
www.incomegambling.com

Random Phenomena

www.incomegambling.com

- Outcome is uncertain
  - In the short-run, outcomes are highly random
  - In the long-run, outcomes are very predictable
- Probability quantifies long-run randomness

Law of Large Numbers



▲ **FIGURE 5.1: The Cumulative Proportion of Times a 6 Occurs, for a Simulation of 100 Rolls of a Fair Die.** The horizontal axis of this MINITAB figure reports the number of the trial, and the vertical axis reports the cumulative proportion of 6s observed by then. **Question:** The first four rolls of the die were 6, 2, 6, and 5. How can you find the cumulative proportion of 6s after each of the first four trials?

As # of trials increase, the proportion of an outcome approaches a particular number (probability): In the long run, 1/6 of die tosses is 6.

Independent Trials

- Trials are *independent* if outcome of one trial is not affected by outcome of another
  - If 20 heads in a row, you are not due a tail - the probability is still 1/2

www.youthnoise.com

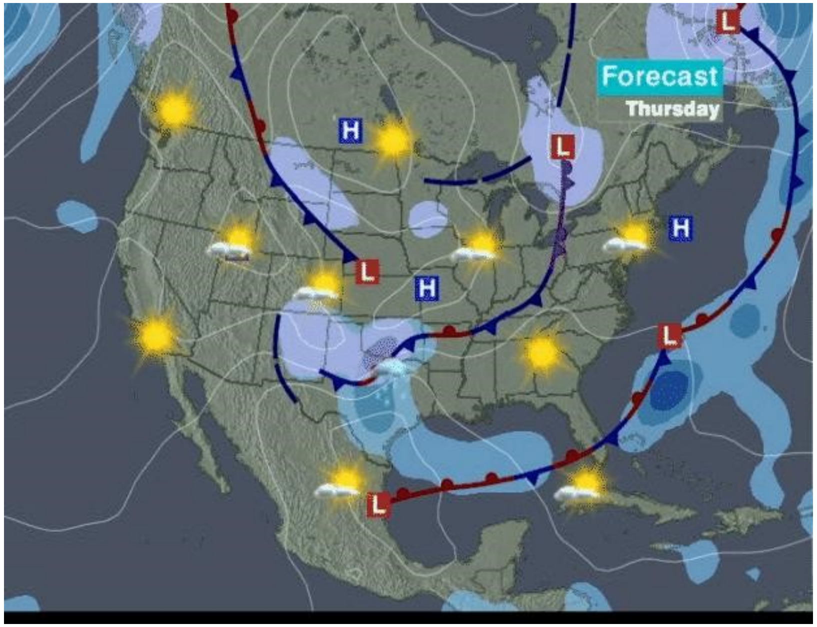
How Can We Find Probabilities?

oumathclub.files.wordpress.com

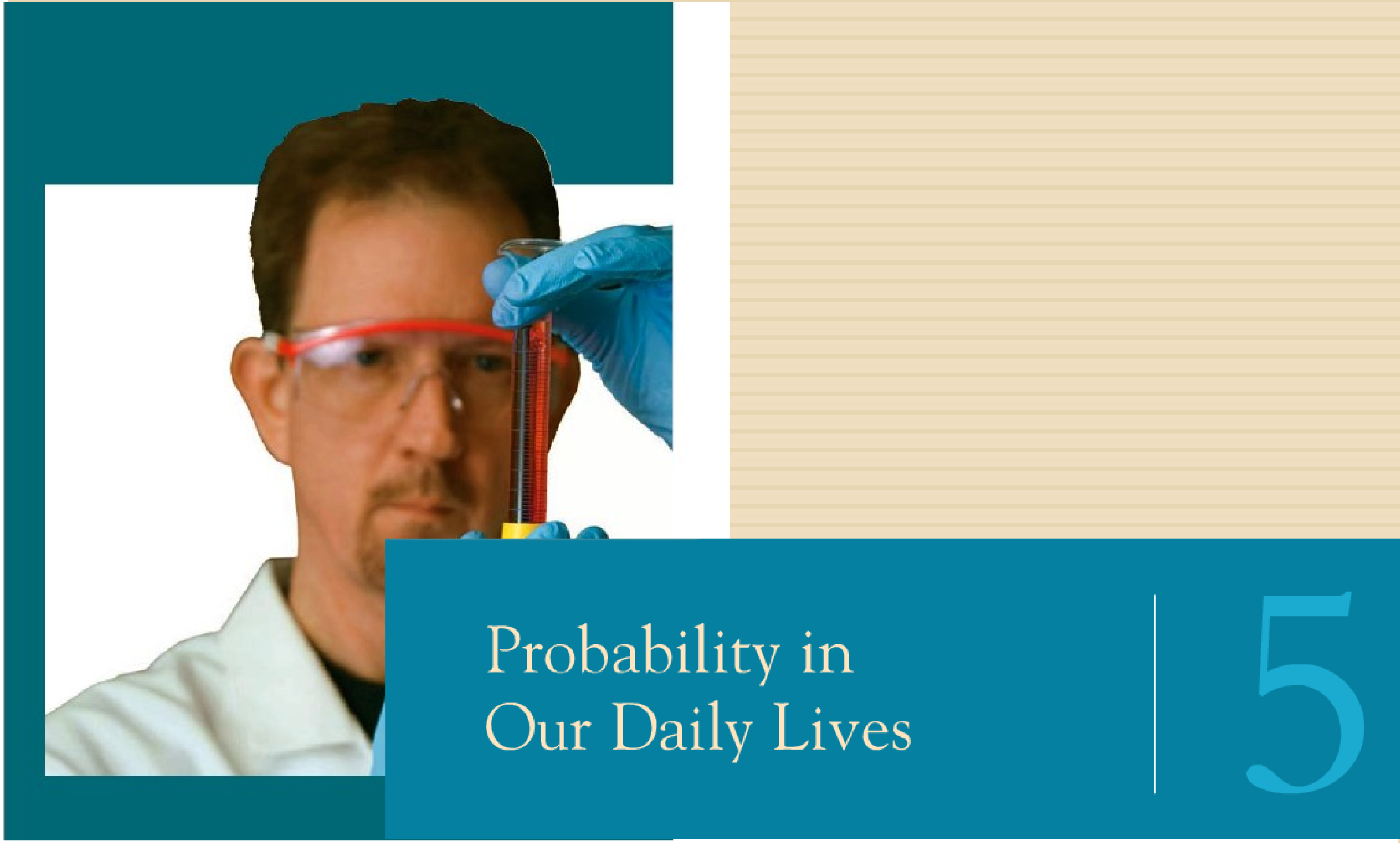
- Theoretical – We assume – some outcomes are equally likely
- Practical – Observe many trials: use sample proportion to estimate probability

Relative Frequency vs. Subjective

- *Relative frequency definition of probability* – proportion outcome occurs in very large number of trials
- *Subjective definition* – degree of belief outcome will occur based on available information
- Bayesian statistics uses subjective probability as its foundation



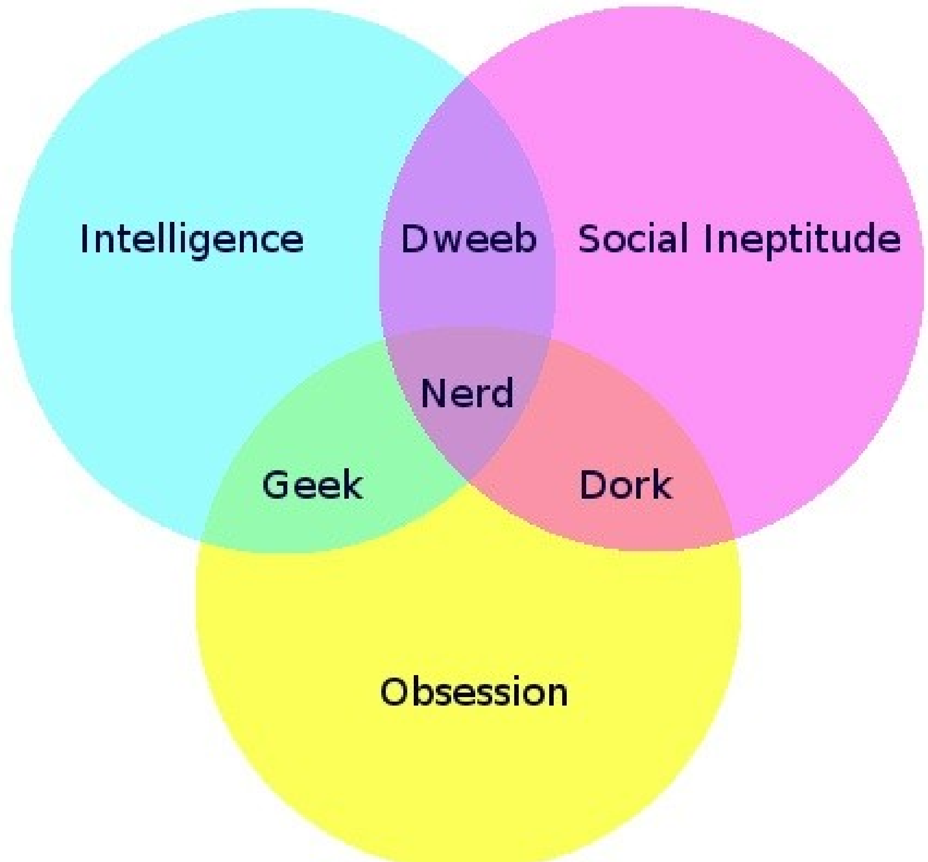
i.cnn.net



5.2 How Can We Find Probabilities?

Learning Objectives

1. Sample Space
2. Event
3. Probabilities for and of
   a. Sample space
   b. Event
   c. Pair of events
   d. Union of two events
   e. Intersection of two events

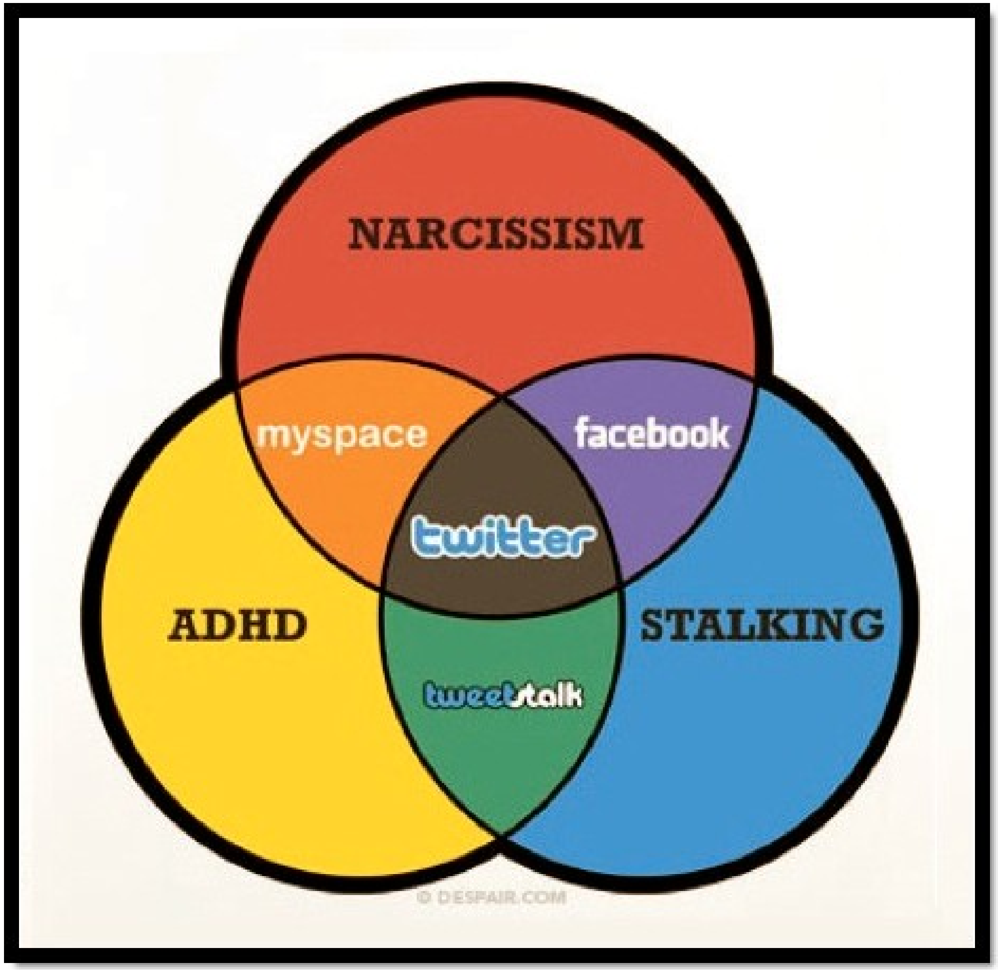


matthewwmason.files.wordpress.com

Sample Space

Sample Space – set of all possible outcomes

Event - Subset of sample space, corresponding to one or a group of outcomes



Probabilities for a Sample Space

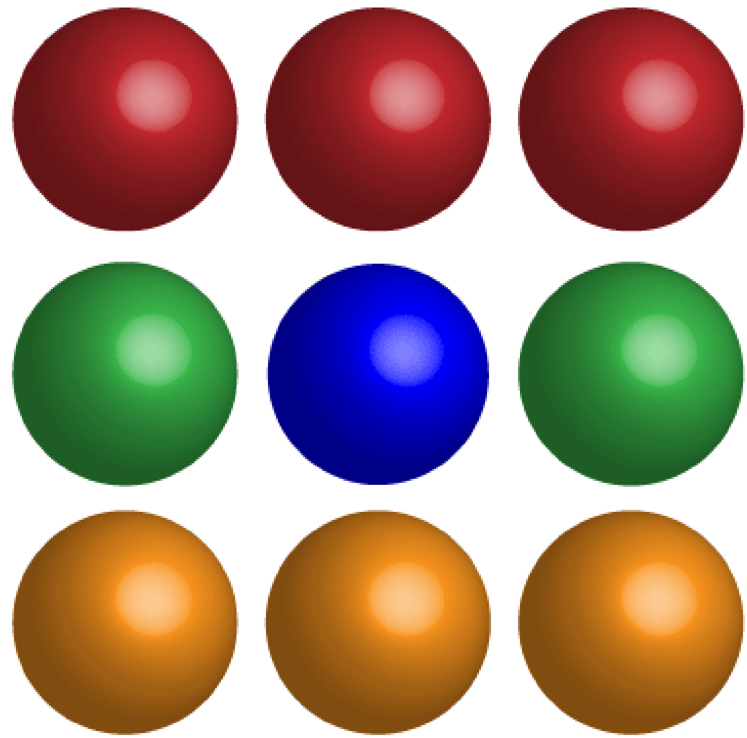
Probability Table	Event:	Probability
Probability of event::	1	0.16
Probability of event::	2	0.22
Probability of event::	3	0.28
Probability of event::	4	0.2
Probability of event::	5	0.14
Probability of event::	0	0
Probability of event::	0	0
Probability of event::	0	0
Probability of event::	0	0
Probability of event::	0	0
Probability of event::	0	0
Probability of event::	0	0
Probability of event::	0	0
Probability of event::	0	0
Probability of event::	0	0
Probability of event::	0	0
Probability of event::	0	0
Probability of event::	0	0
Probability of event::	0	0
Probability of event::	0	0
Probability of event::	0	0

Each outcome has a probability

- All probability is between 0 and 1
- Sum of all probabilities is 1

Probability of an Event

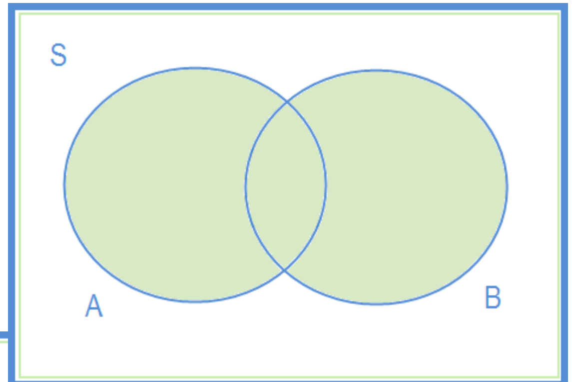
- Probability of event A, P(A), is sum of outcome probabilities in event
- When all outcomes are equally likely:



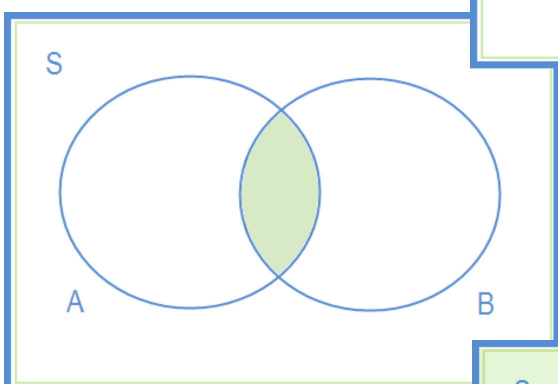
www.georgiamath1234.com

$$P(A) = \frac{\#\text{ outcomes in A}}{\#\text{ outcomes in Sample Space}}$$

Probability Rules of Pairs of Events

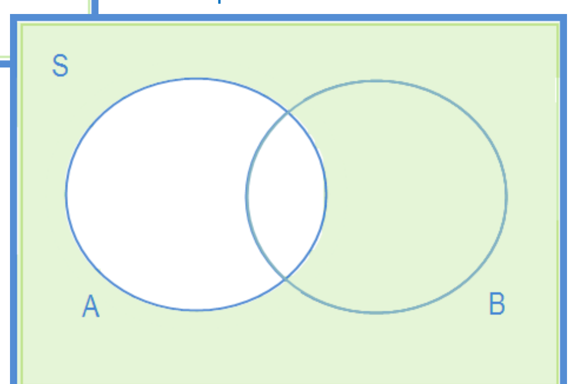


Union of A and B: A or B



Intersection of A and B: A & B

Complement of A: A^C



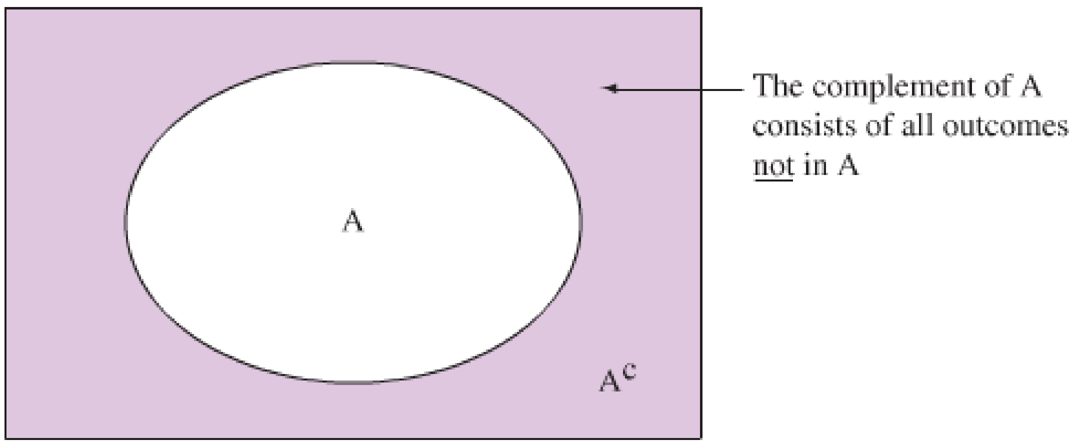
- In either or both events (**Union**)
- In both events (**Intersection**)
- Not in the event (**Complement**)

Complement of Event A

All outcomes in sample space not in A

$P(A) + P(A^c) = 1$

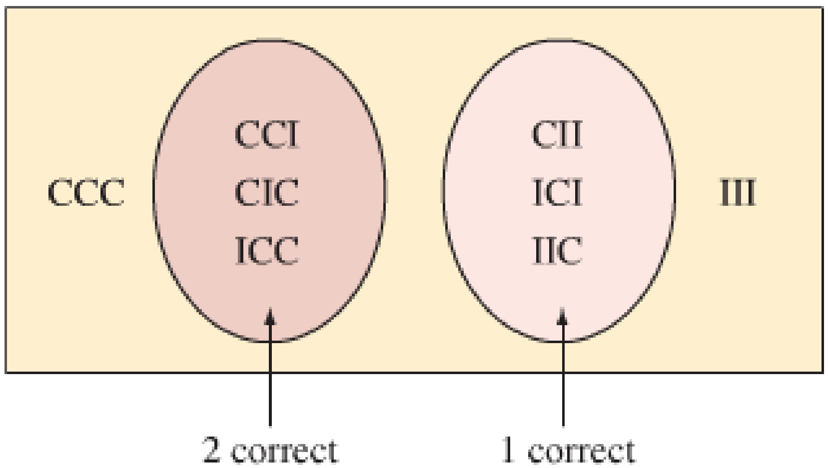
So, $P(A^c) = 1 - P(A)$



▲ **FIGURE 5.4: Venn Diagram Illustrating an Event A and Its Complement A^c.**
Question: Can you sketch a Venn diagram of two events A and B such that they share some common outcomes, but some outcomes are not in A or in B?

Disjoint Events

A and B are disjoint if they have no common outcomes

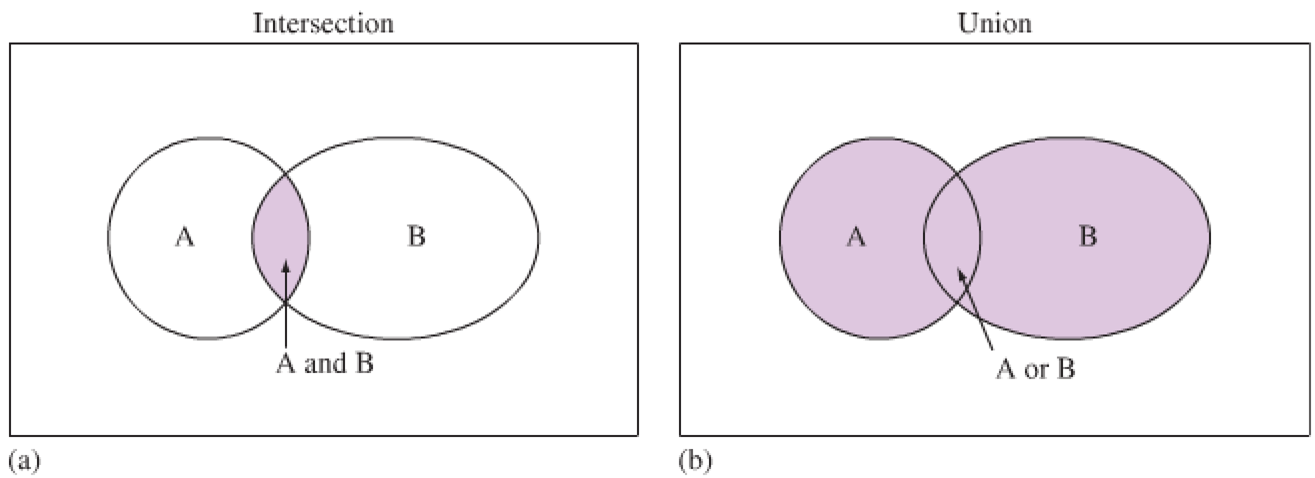


▲ **FIGURE 5.5: Venn Diagram Illustrating Disjoint Events.** The event of a student answering exactly one question correctly is disjoint from the event of answering exactly two questions correctly. **Question:** Identify on this figure the event that the student answers the first question correctly. Is this event disjoint from either of the two events identified in the Venn diagram?

Intersection & Union of A and B

Intersection is all outcomes in both A and B

Union is all outcomes in either A or B or both



▲ **FIGURE 5.6: The Intersection and the Union of Two Events.** Intersection means A occurs *and* B occurs, denoted "A and B." The intersection consists of the shaded "overlap" part in Figure 5.6(a). Union means A occurs *or* B occurs *or* both occur, denoted "A or B." It consists of all the shaded parts in Figure 5.6(b). **Question:** How could you find P(A or B) if you know P(A), P(B), and P(A and B)?

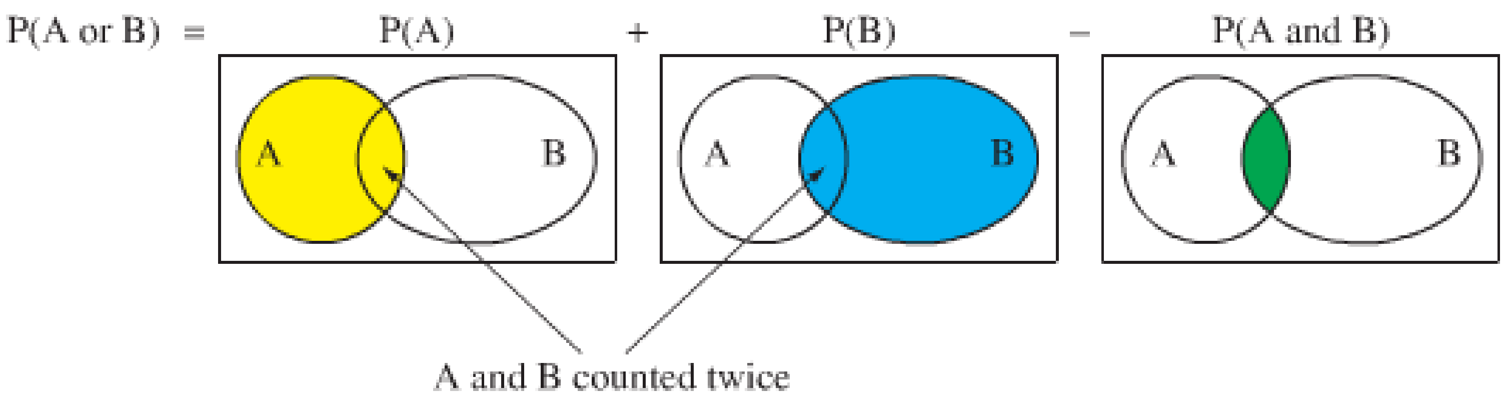
Probability of Union of Two Events

Addition Rule for the *union* of any two events,

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

If *disjoint*, P(A and B) = 0, so

$$P(A \text{ or } B) = P(A) + P(B)$$



▲ **FIGURE 5.7: The Probability of the Union, Outcomes in A or B or Both.** Add P(A) to P(B) and subtract P(A and B) to adjust for outcomes counted twice. **Question:** When does P(A or B) = P(A) + P(B)?

Chances of Being Audited?

P(A and B) = ? when A is being audited and B is income greater than $100,000

	Audited		
Income Level	Yes	No	**Total**
Under $25,000	90	14010	**14100**
$25,000–$49,999	71	30629	**30700**
$50,000–$99,999	69	24631	**24700**
$100,000 or more	80	10620	**10700**
Total	**310**	**79890**	**80200**

Source: *Statistical Abstract of the United States: 2003.*

Probability of Intersection of Two Events

Multiplication Rule for intersection of two independent events, A and B,

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

IN WORDS

Independent trials means that what happens on one trial is not influenced by what happens on the other trial.

	Audited		
Income Level	Yes	No	**Total**
Under $25,000	90	14010	**14100**
$25,000–$49,999	71	30629	**30700**
$50,000–$99,999	69	24631	**24700**
$100,000 or more	80	10620	**10700**
Total	**310**	**79890**	**80200**

Source: Statistical Abstract of the United States: 2003.

Assuming Independence

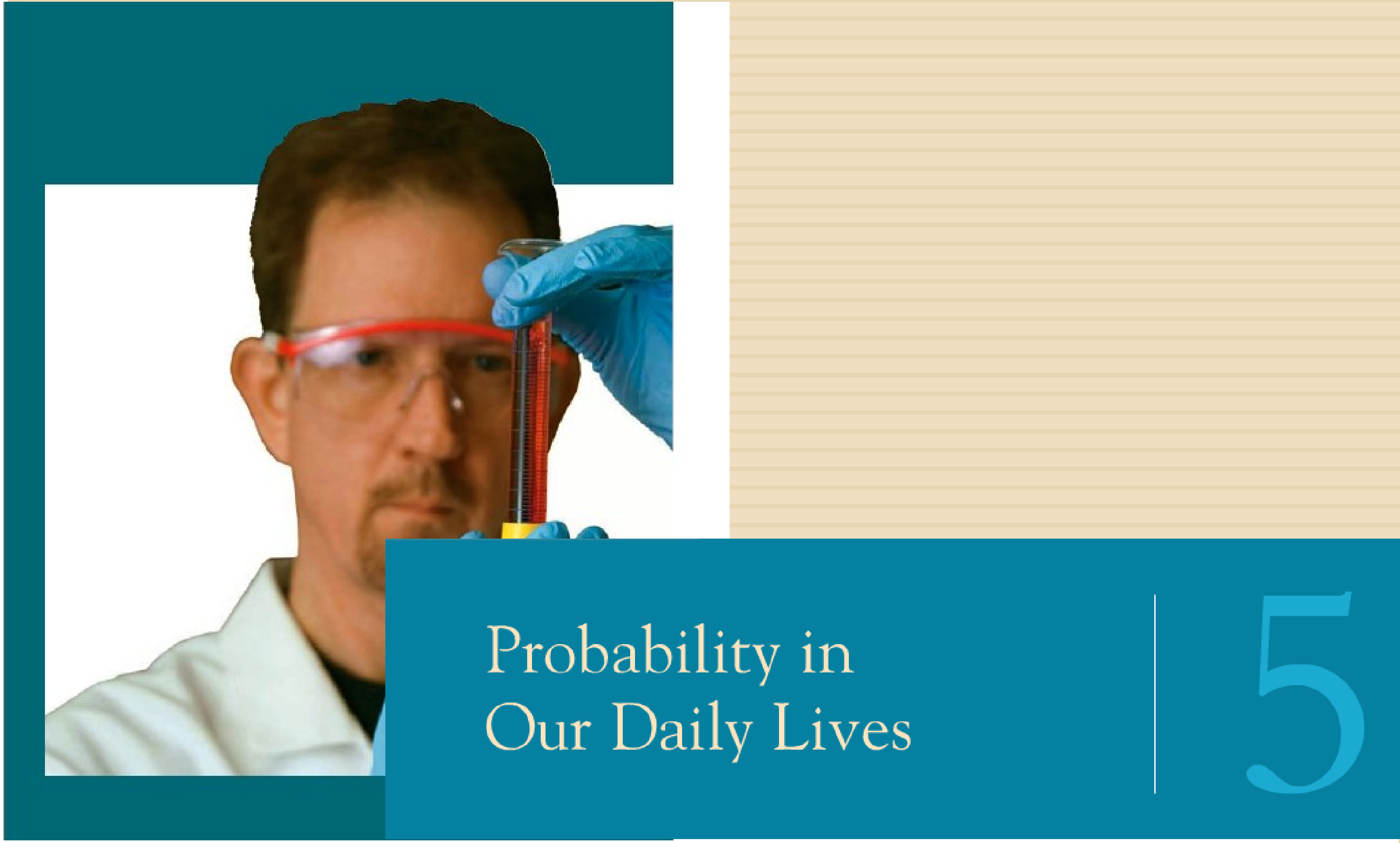
	2nd Question	
1st Question	C	I
C	0.58	0.05
I	0.11	0.26
	A and B	

A = 1st Question Correct

B = 2nd Question Correct

www.youthcentral.vic.gov.au

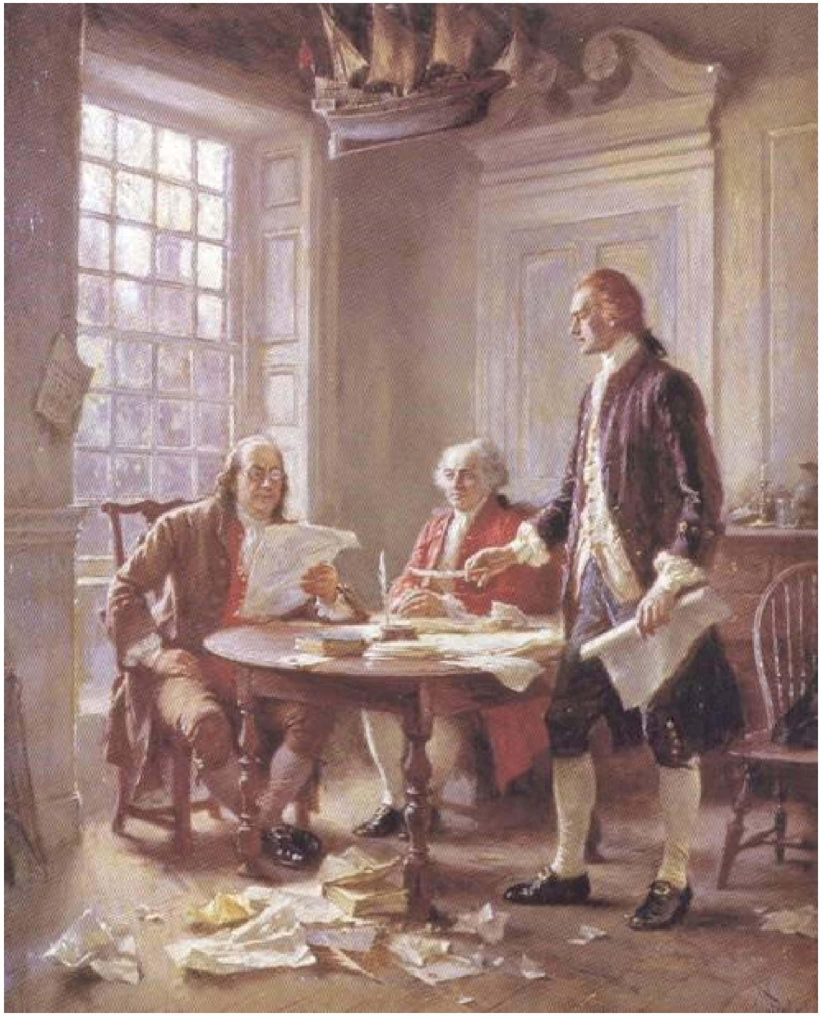
Don't assume independence!



5.3 Conditional Probability: Probability of A, Given B

Learning Objectives

1. Conditional probability
2. General multiplication rule for P(A and B)
3. Independent events defined using conditional probability



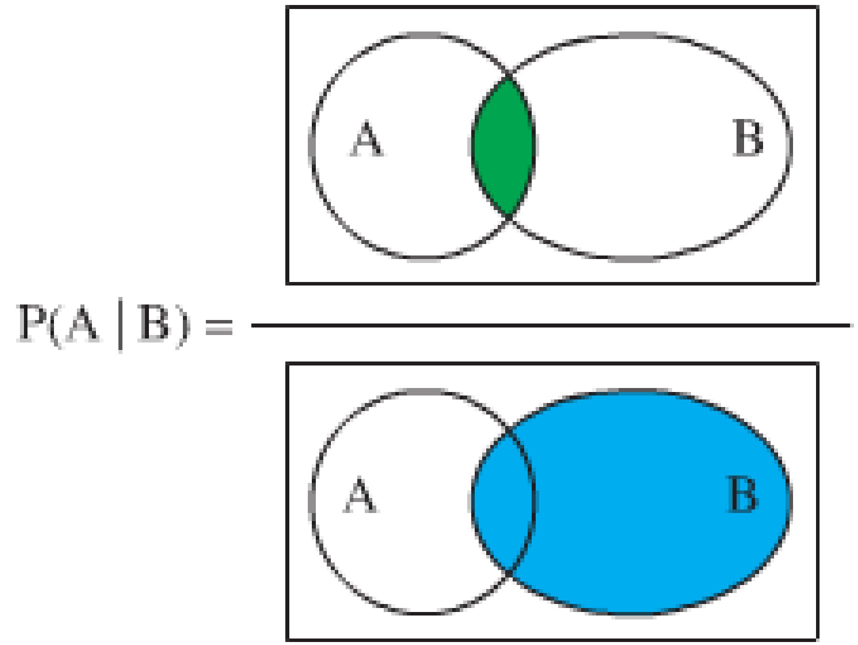
dnewell1.files.wordpress.com

Conditional Probability

- Probability of event A, given that event B has occurred, is:

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

The vertical slash represent the word "given". Of the times that B occurs, P(A|B) is the proportion of times that A also occurs



Probability Distribution

TABLE 5.3: Probabilities of Taxpayers at the Eight Possible Combinations of Income and Whether Audited

Each frequency in Table 5.2 was divided by 80200 to obtain the cell probabilities shown here, such as 90/80200 = 0.0011.

	Whether Audited		
Income Level	Yes	No	**Total**
	These 8 probabilities sum to 1.0		
Under $25,000	0.0011	0.1747	**0.1758**
$25,000–$49,999	0.0009	0.3819	**0.3828**
$50,000–$99,999	0.0009	0.3071	**0.3080**
≥$100,000	0.0010	0.1324	**0.1334**
Total	**0.0039**	**0.9961**	**1.00**

	Audited		
Income Level	Yes	No	**Total**
Under $25,000	90	14010	**14100**
$25,000–$49,999	71	30629	**30700**
$50,000–$99,999	69	24631	**24700**
$100,000 or more	80	10620	**10700**
Total	**310**	**79890**	**80200**

Tax Forms by Income and Audit Status

Probability of audit (event A), given income ≥ $100,000 (event B)?

$$P(A \mid B) = \frac{P(A \text{ and } B)}{P(B)} = \frac{0.0010}{0.1334} = 0.007$$

Income Level	Yes	No	Total
Under $25,000	0.0011	0.1747	**0.1758**
$25,000–$49,999	0.0009	0.3819	**0.3828**
$50,000–$99,999	0.0009	0.3071	**0.3080**
≥$100,000	0.0010	0.1324	**0.1334**
Total	**0.0039**	**0.9961**	**1.0000**

www.csbgl.org

Multiplication Rule for P(A and B)

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

and

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

Independent Events Defined Using Conditional Probabilities

Events A and B are *independent* if the probability one occurs is unaffected by whether other occurs

Dependence Tests:

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

www.psychologytoday.com

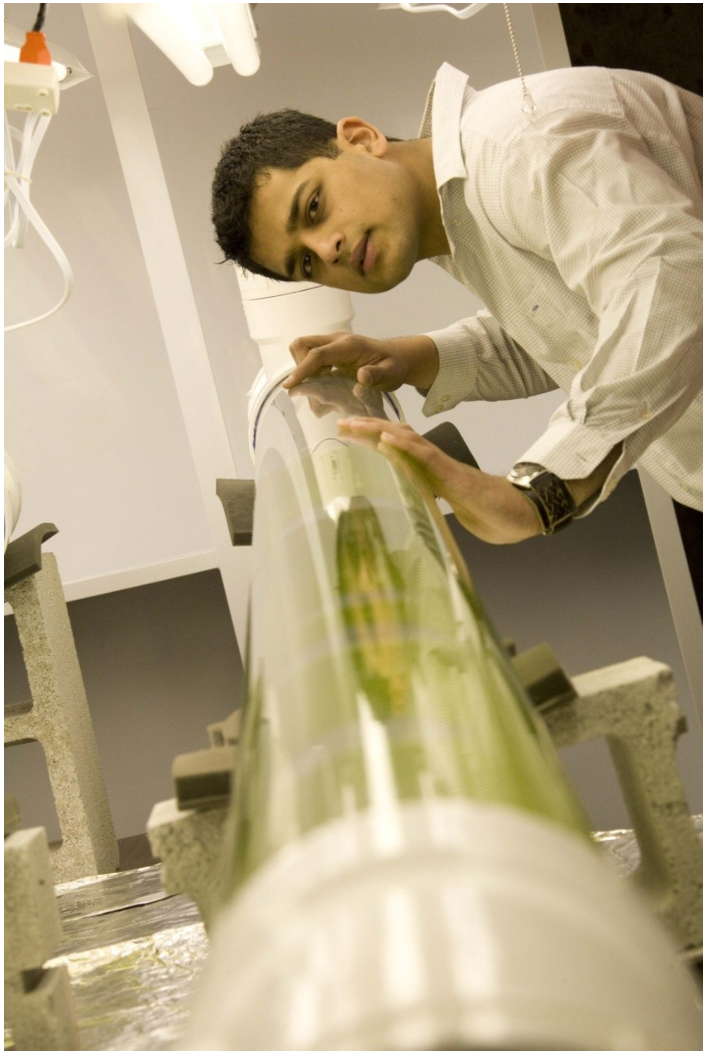
Probability in
Our Daily Lives

5

5.4 Applying the Probability Rules

Learning Objectives

1. Is a “Coincidence” Truly an Unusual Event?
2. Probability Model
3. Probabilities and Diagnostic Testing
4. Simulation



Is Coincidence Truly Unusual Event?

Stonehenge
blog.silive.com

Law of Very Large Numbers – if something has very large number of opportunities to happen, occasionally it will happen, even if highly unusual

Probability Model

Unlike examples, deciding whether equally likely or independent is difficult

- We must specify a *probability model* that spells out assumptions

Probability Model

A **probability model** specifies the possible outcomes for a sample space and provides assumptions on which the probability calculations for events composed of those outcomes are based.

Probability Model

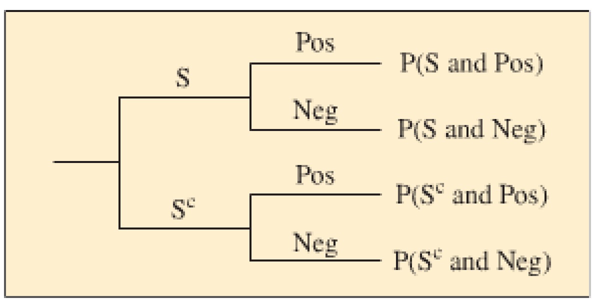
farm4.static.flickr.com

- **Sensitivity** = P(POS | S)
- **Specificity** = P(NEG | S^C)

TABLE 5.6: Probabilities of Correct and Incorrect Results in Diagnostic Testing

The probabilities in the body of the table refer to the test result, conditional on whether the state (S) is truly present. The sensitivity and specificity are the probabilities of the two types of correct diagnoses.

	DIAGNOSTIC TEST RESULT		
State Present?	**Positive (POS)**	**Negative (NEG)**	**Total Probability**
Yes (S)	Sensitivity P(POS | S)	False negative rate P(NEG | S)	1.0
No (S^c)	False positive rate P(POS | S^c)	Specificity P(NEG | S^c)	1.0



Simulation

assets1.csc.com

Simulation is used for probabilities difficult to find with ordinary reasoning

1. Identify random phenomenon
2. Describe how to simulate
3. Repeat (at least 1000)
4. Summarize and state conclusion

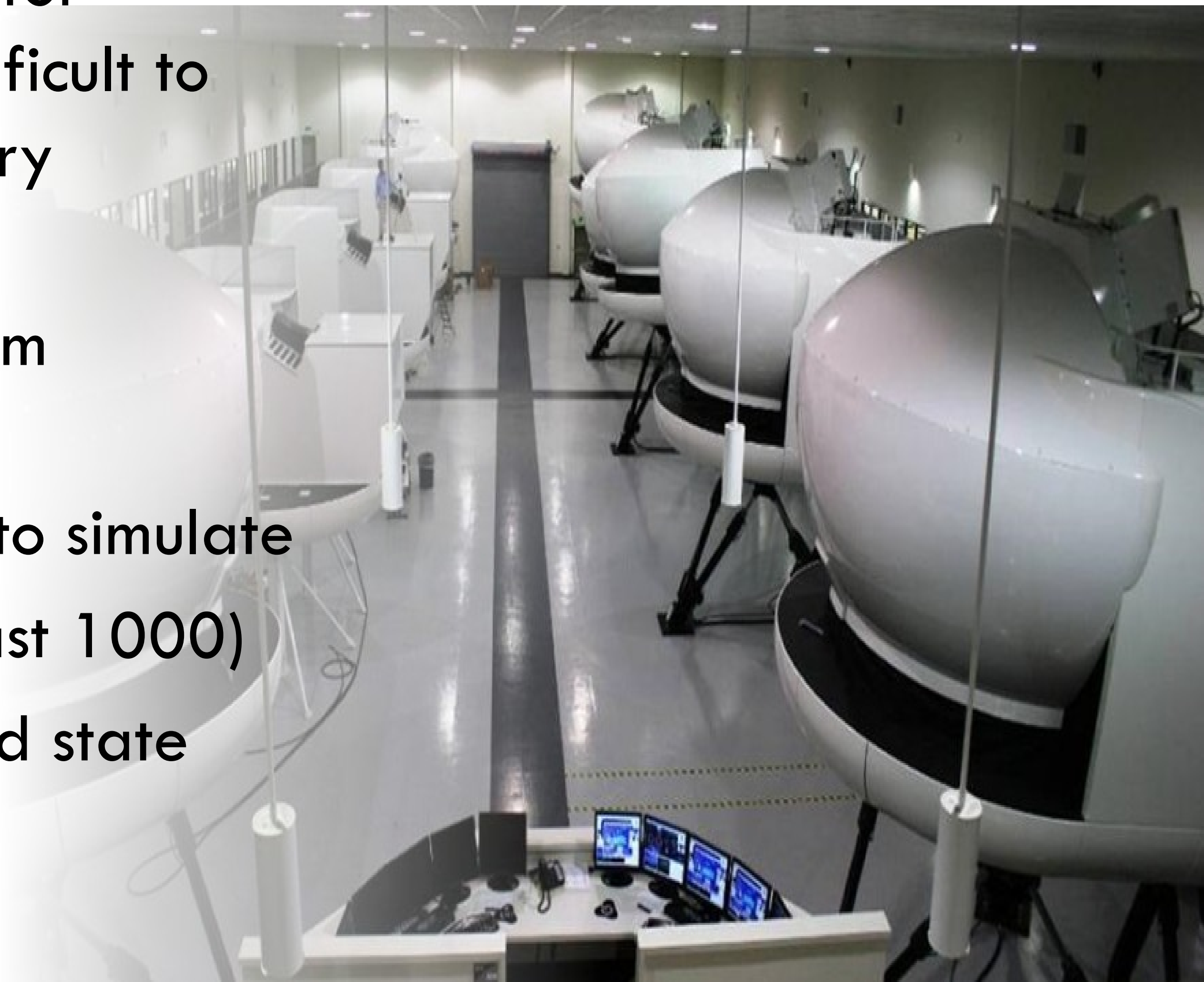


Image Sources

Statistics: The Art and Science of Learning from Data, 2nd Edition, Agresti and Franklin

http://www.incomegambling.com/img/upload/image______________________________.jpg

http://www.limeduck.com/wp-content/uploads/2008/09/latteart.jpg

http://www.youthnoise.com/Assets/Image/Gill%27s%20images/coin_flip.jpg

http://oumathclub.files.wordpress.com/2009/11/52589.jpg

http://i.cnn.net/cnn/.element/img/1.0/weather/cnn/namerica_forecast.jpg

http://matthewwmason.files.wordpress.com/2009/09/nerd-venn-diagram-20090915-092804.jpg

http://kk.org/ct2/Socialven.jpg

http://tredway.files.wordpress.com/2009/04/test_taking1.jpg

http://www.stathomeworkhelper.com/freqdist%20prob.JPG

http://www.georgiamath1234.com/math_images/nine_marbles.gif

http://www.youthcentral.vic.gov.au/digitalAssets/30975_StudyTips.jpg

http://dnewell1.files.wordpress.com/2008/07/declaration_of_independence.jpg

http://www.csbgl.org/images/Girls_Audit2.jpg

http://www.psychologytoday.com/files/u114/dice_home.jpg

http://blog.silive.com/weather/2008/09/stonehenge-wallpaper-4.jpg

http://farm4.static.flickr.com/3126/2589960921_59590ba85a.jpg