CLASS #5

Probability, 1016-351

Rochester Institute of Technology

Sections 2.1 and 2.2 of the textbook

Probability

Methods of assigning probability

1. relative frequency
   from observations

   example: flip a coin many times
   and take the fraction of heads
   as the probability of a head

These are called empirical probabilities.

Probability

Methods of assigning probability continued

2. <u>relative frequency</u>

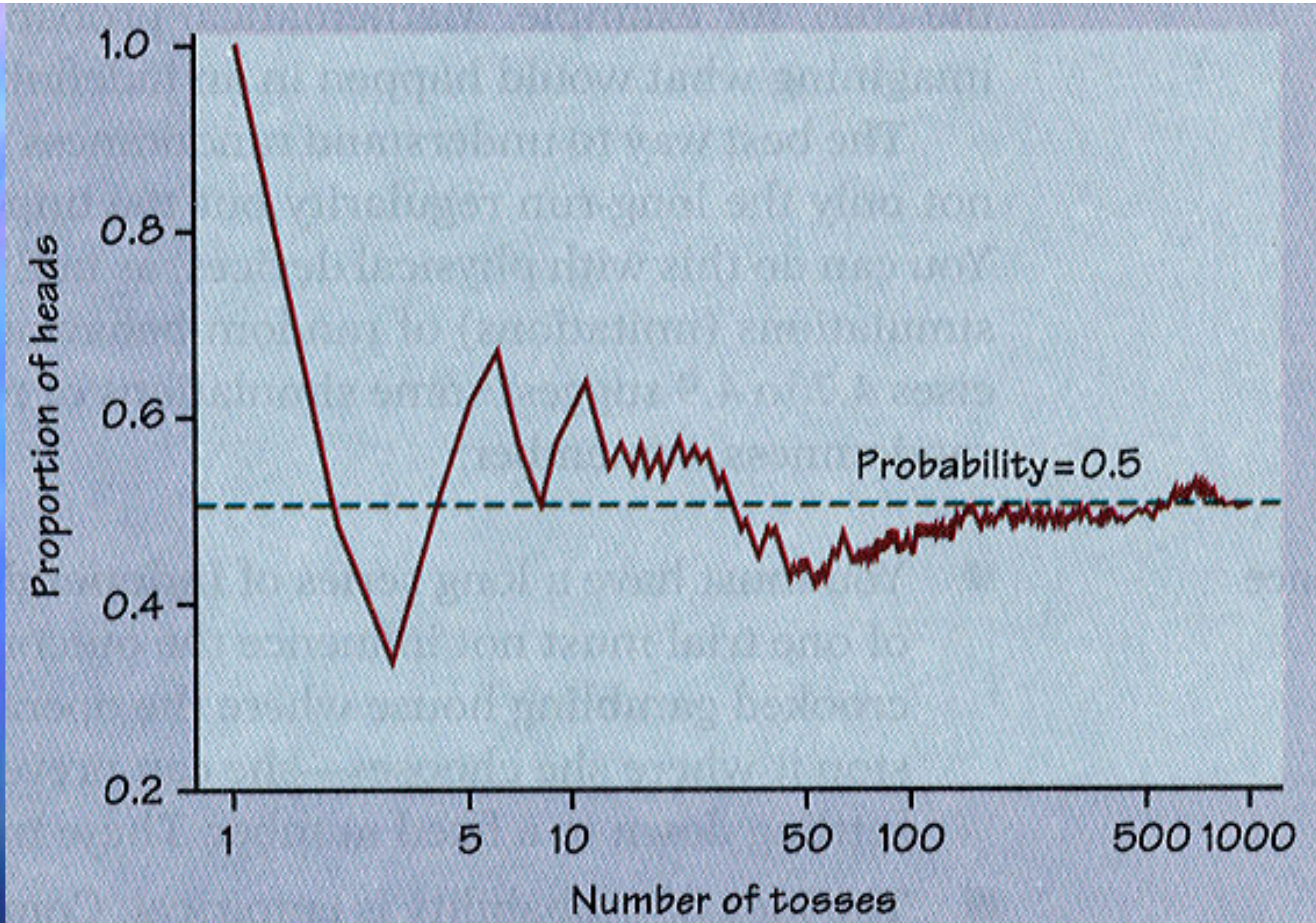
   <u>from physical setup</u>

   example: examine a coin and decide that it is fair so that the probability of a head is 1/2

Probability

Methods of assigning probability continued

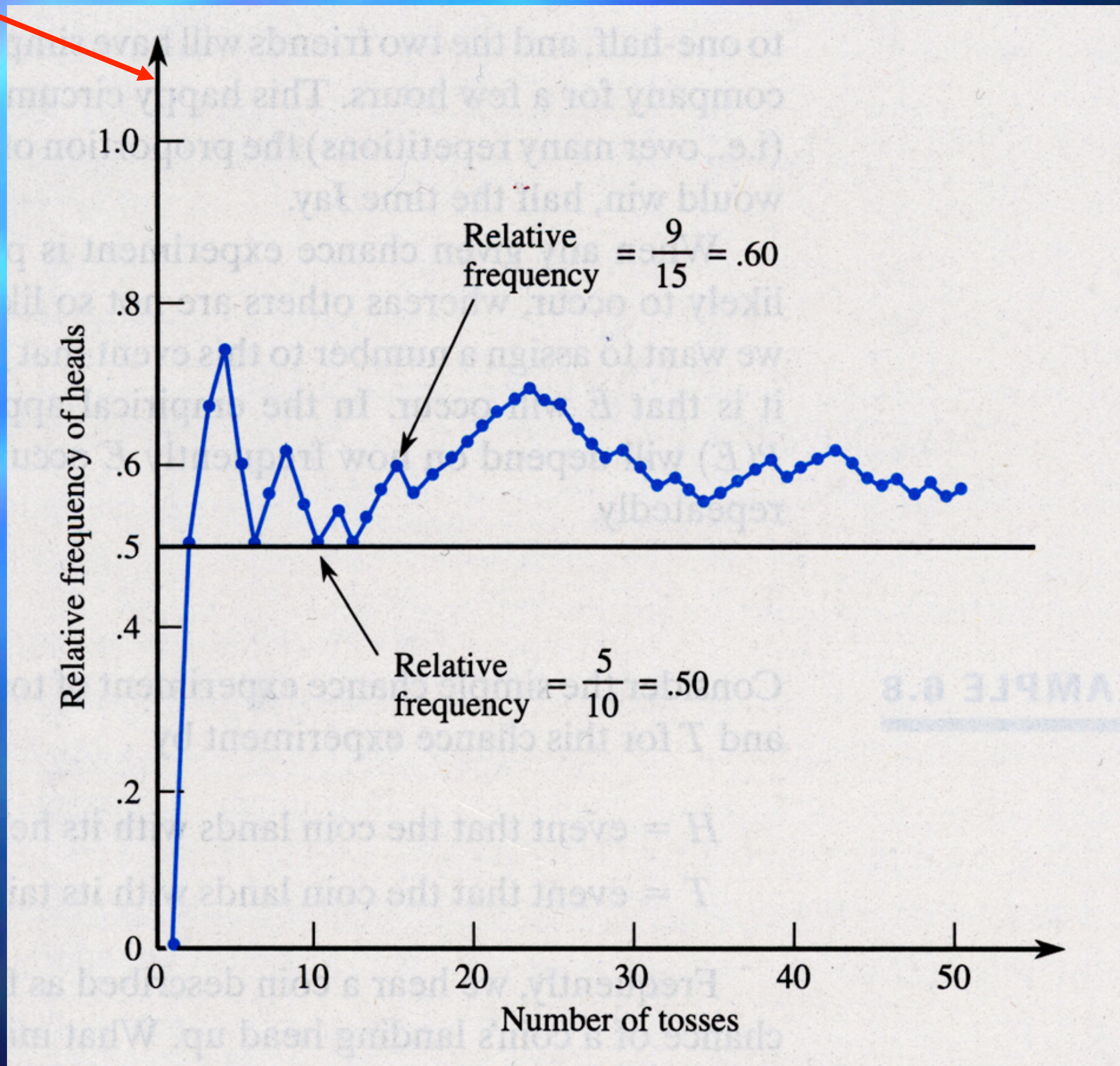
3. limiting relative frequency (long-run relative frequency)
   called limiting because we take the number of trials to the limit (a very large number)



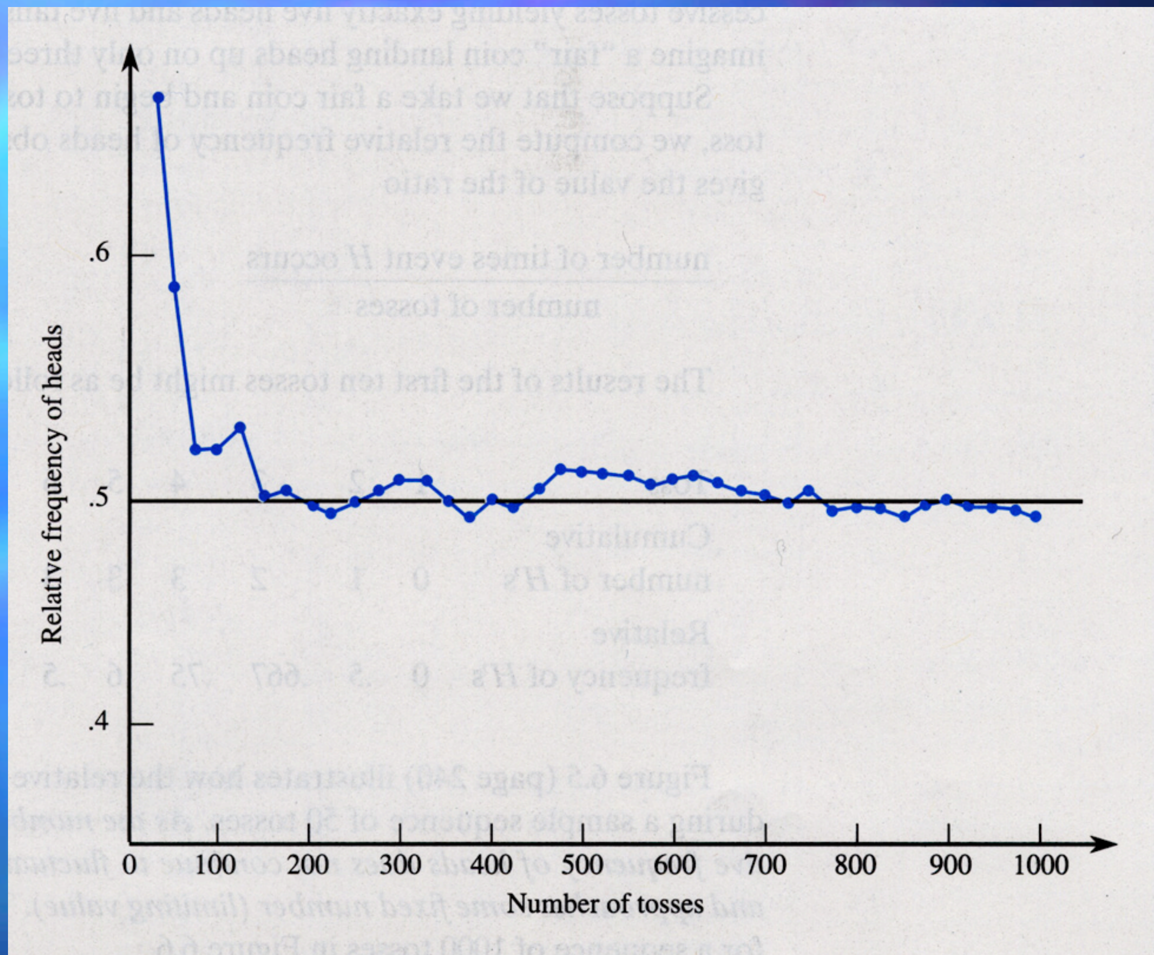
long-run relative frequency

long-run relative frequency

use cumulative relative frequencies



long-run relative frequency



This is an example of a “law of large numbers.”

Probability

Methods of assigning probability continued

4. <u>subjective</u> <u>assignment</u>

example: The probability that the Buffalo Bills will win the Super Bowl this January is 0.8.

These probabilities are is a kind of measure of belief.

Probability

Methods of assigning probability continued

5. any mathematical formulation satisfying the "rules" of probability

   If it looks like a probability and acts like a probability, then it is a probability.

Example of relative frequencies from a physical setup:

the number of heads obtained when flipping a fair coin three times

Elements of the sample space	x	frequency	f(x)
t,t,t	0	1	1/8
h,t,t; t,h,t; t,t,h	1	3	3/8
h,h,t; h,t,h; t,h,h	2	3	3/8
h,h,h	3	1	1/8
	sums	8	1

x is the number of heads.

f(1) = 3/8

sample space *S*

outcome or event (often labeled A and B)

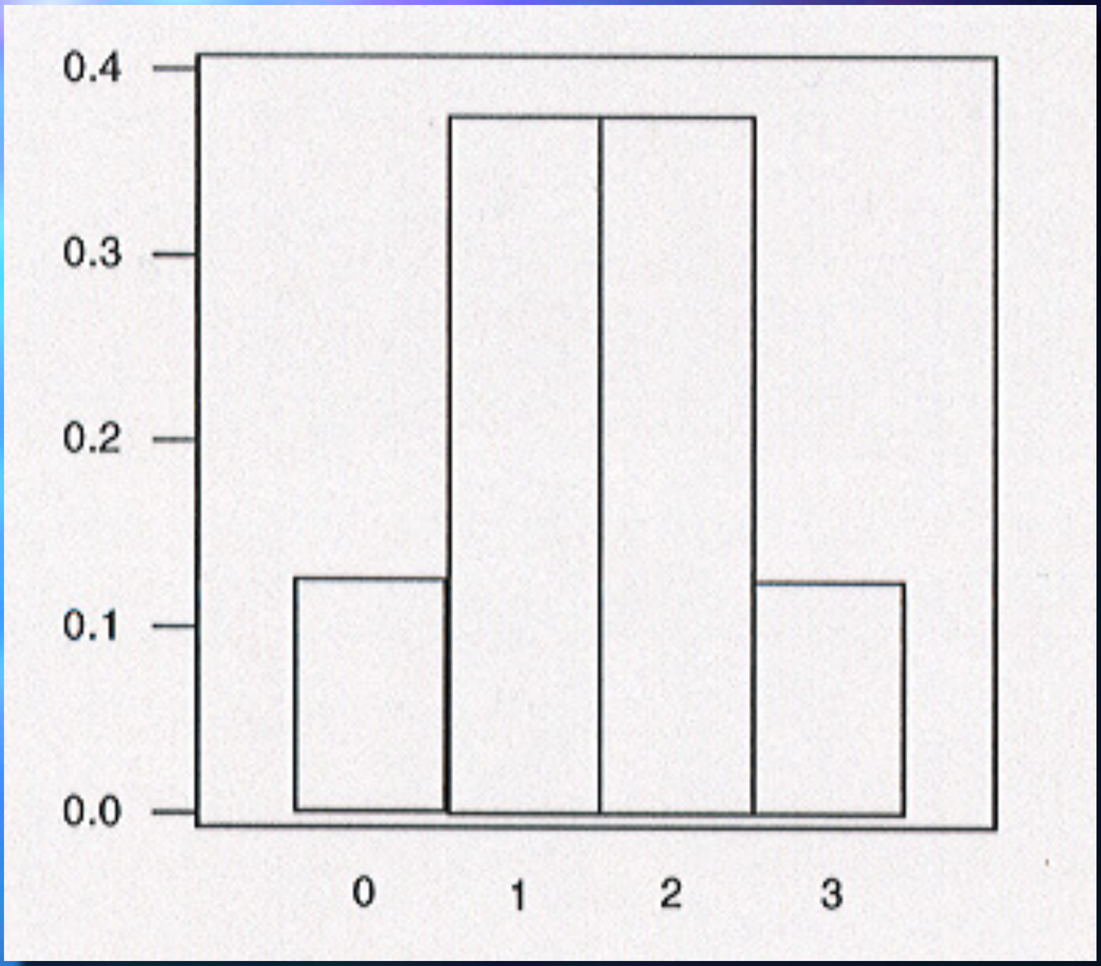
- simple versus compound
- element

random variable

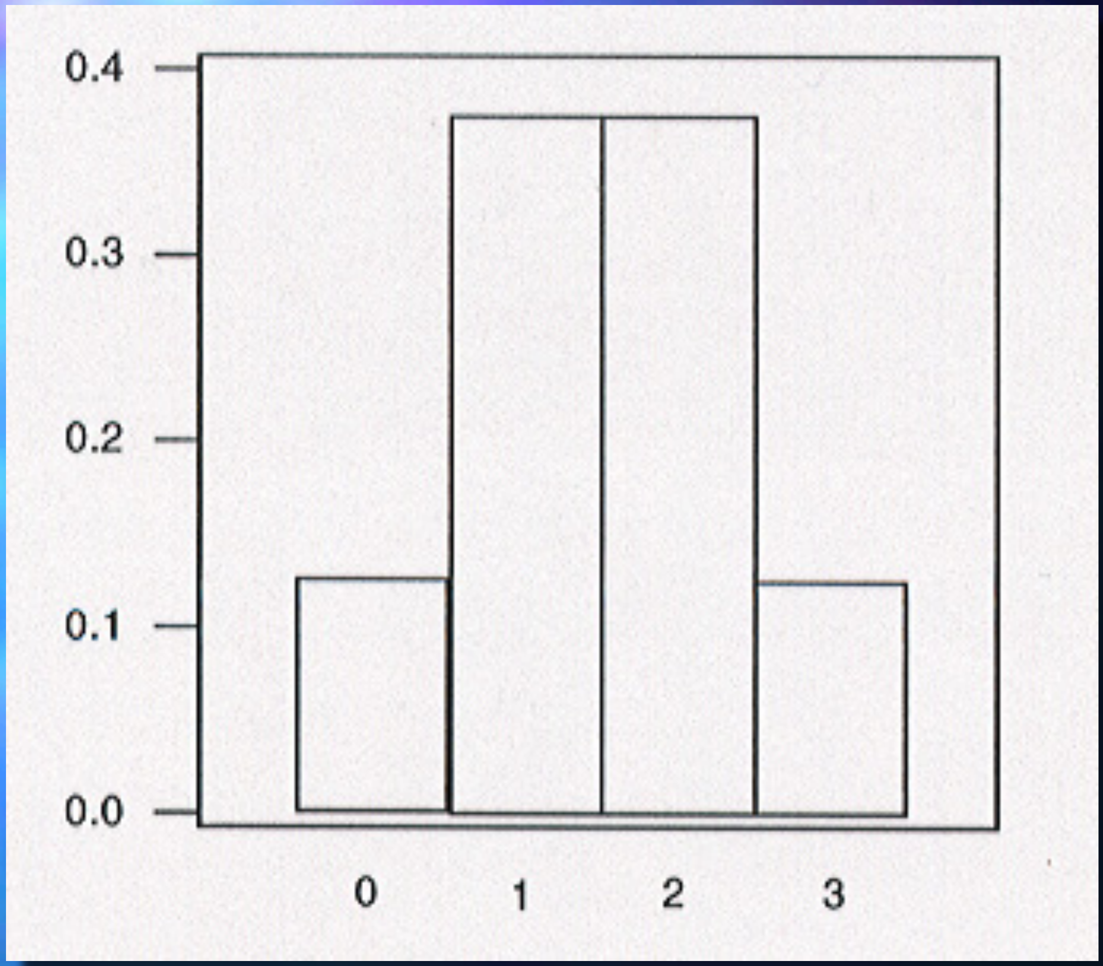
- discrete versus continuous

probability histogram

sampling distribution



probability histogram for the number of heads in three tosses



area in the rectangles = (1)(1/8) + (1)(3/8)
+ (1)(3/8) + (1)(1/8) = 1

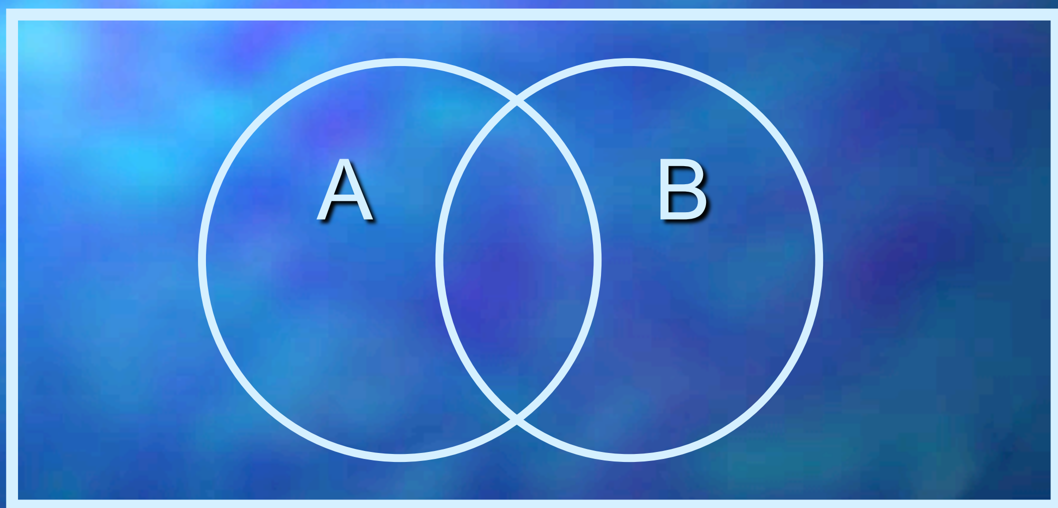
There are three issues.

1. We need to review sets. (That is Section 2.1.)
2. We need the rules of probability. (That is Section 2.2.)
3. We need to be able to count the number of ways an event can happen, so that we can find the fraction of the time that the event occurs. (That is Section 2.3.)

Probability is used to think about events that are between certainty of occurring and chaos.

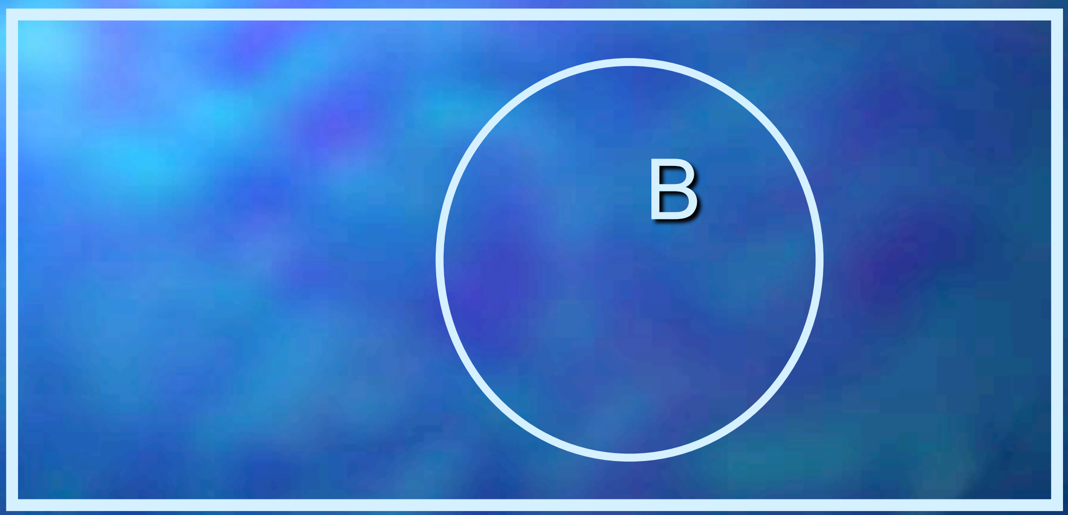
Venn diagrams

John Venn (1834-1923), a British logician



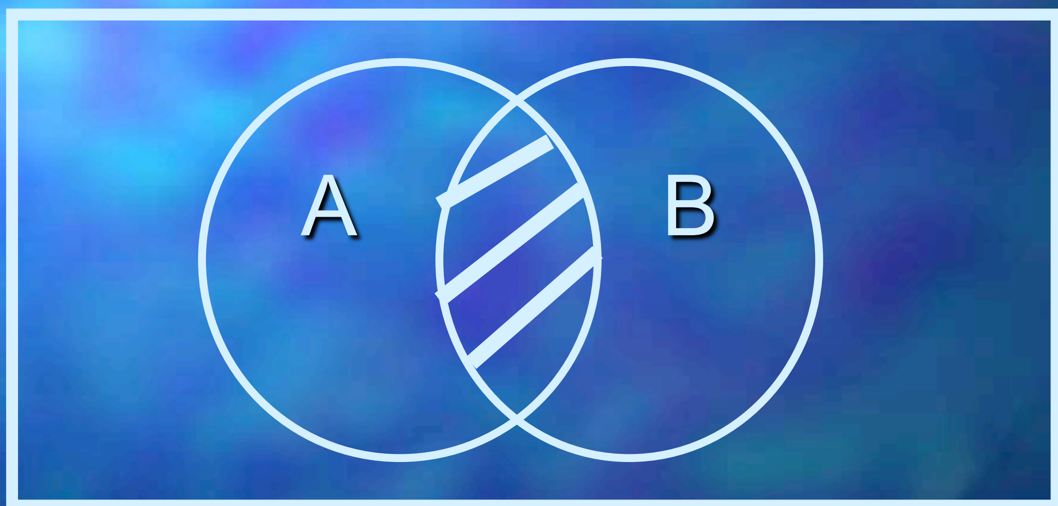
Venn diagrams

Venn diagrams



Venn diagrams

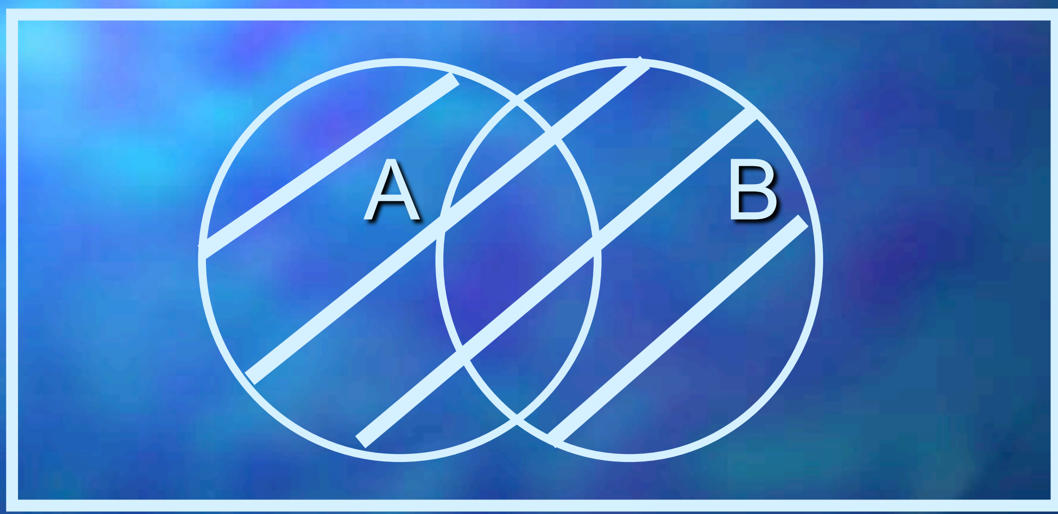
A and B



intersection of A and B; A ∏ B

Venn diagrams

A or B



A union B; A U B

Venn diagrams

not A

complement of A; A^c

Venn diagrams

A and B are disjoint events (no “overlap”)

Also known as mutually exclusive events

Axioms of Probability page 51

Axiom 1: For any event A, $P(A) \geq 0$.

Think of probability as the fraction of the time that something (A) happens or as a relative frequency. Combined with Axiom 2, we have $0 \leq P(A) \leq 1$.

Axioms of Probability

Axiom 2: $P(S) = 1$

Axiom 3: If $A_1, A_2, \ldots, A_k$ is a collection of mutually exclusive events, then $P(A_1 \cup A_2 \cup \ldots \cup A_k) = \Sigma P(A_i)$ where the sum is over $i = 1$ to $i = k$ or to $i = \infty$.

Example

P(A or B) = P(A) + P(B)

for A and B disjoint events

example: in the coin tossing problem from a previous slide (and the next slide) for three tosses of a fair coin,

P(2 or 3 heads) = P(2 heads) + P(3 heads) =

3/8 + 1/8 = 4/8 = 1/2

Think of “3 heads” as “ exactly 3 heads.”

Elements of the sample space	x	frequency	f(x)
t,t,t	0	1	1/8
h,t,t; t,h,t; t,t,h	1	3	3/8
h,h,t; h,t,h; t,h,h	2	3	3/8
h,h,h	3	1	1/8
	sums	8	1

x is the number of heads.

Properties of Probability

Proposition on page 54: For any event A,
$P(A) = 1 - P(A^c)$

This is just a rewrite of $1 = P(S) = P(A \cup A^c) = P(A) + P(A^c)$, where the last equality is justified by Axiom 3.

Example

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

We know P(no rain in London tomorrow) = 0.3.

P(rain in London tomorrow)

= 1 - P(no rain in London tomorrow)

= 1 - 0.3

= 0.7

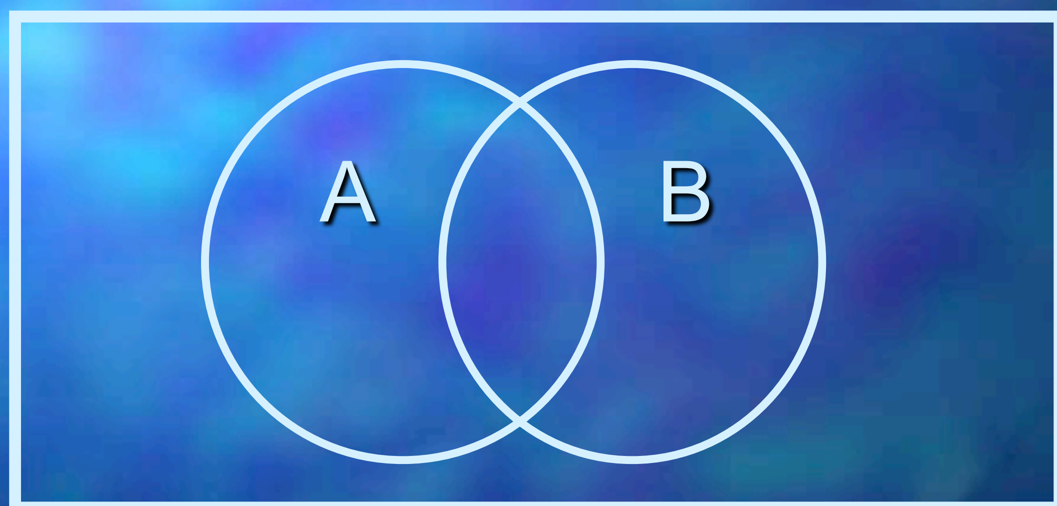
Properties of Probability

Page 55: If A and B are mutually exclusive, then $P(A \prod B) = 0$.

Since $A \prod B$ is empty, $(A \prod B)^c = S$. So, $P(A \prod B) = 1 - P((A \prod B)^c) = 1 - P(S)$ $= 1 - 1 = 0$.

Proposition on page 55 For any two events A and B,

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

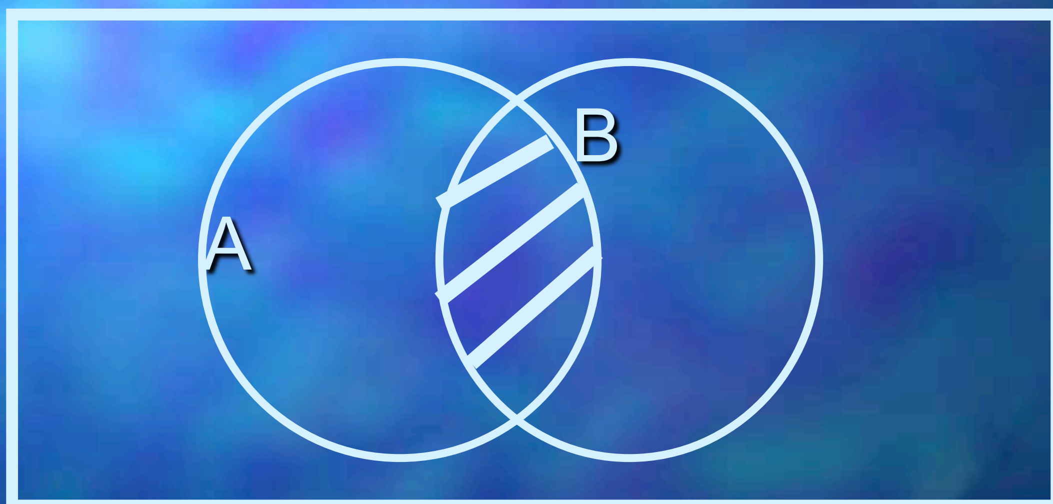


$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$

A or B

P(A or B) = P(A) + P(B) - P(A and B)

A and B



Also, see the proof on page 55.

Example

Given

P(A) = P(female) = 0.320

P(B) = P(engineering major) = 0.071

P(A and B) = P(both female and engineering major) = 0.005

P(A or B) = P(A) + P(B) - P(A and B)

P(E or F)

= P(engineering major or female)

= 0.071 + 0.320 – 0.005

= 0.391 – 0.005 = 0.386

According to the article “Greek Perceptions at RIT” by Monica Donovan on pages 16-19 of the April 29, 2005, issue of Reporter, 31% of RIT students were female.

According to the article “Women Courted by Technology” on pages 18-21 of the April 6, 2007 issue of Reporter, 32% of RIT students were female in Fall 2006.

According to the article “is there reverse discrimination at rit?” by Veena Chatti on pages 20-21 of the November 2, 2007, issue of Reporter, 4,974 RIT students were female and 10,583 RIT students were male. So,

$$[4974/(4974 + 10583)]100 = 31.972...\%$$

or about 32.0% were female.

According to the article “Women in Colleges at RIT?” by Justin Claire on page 11 of the November 6, 2009, issue of Reporter, in 2008, 4,232 RIT students were female and 8,703 RIT students were male. So,

$$[4232/(4232 + 8703)]100 = 32.71...\%$$

or about 33% were female.

(That article seems to have been about undergraduate students only, but did not say so.)

END OF SLIDES FOR CLASS #5