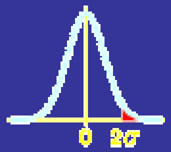
Introduction to Probability

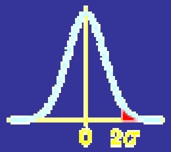
SOCY601—Alan Neustadtl

Introduction to Probability

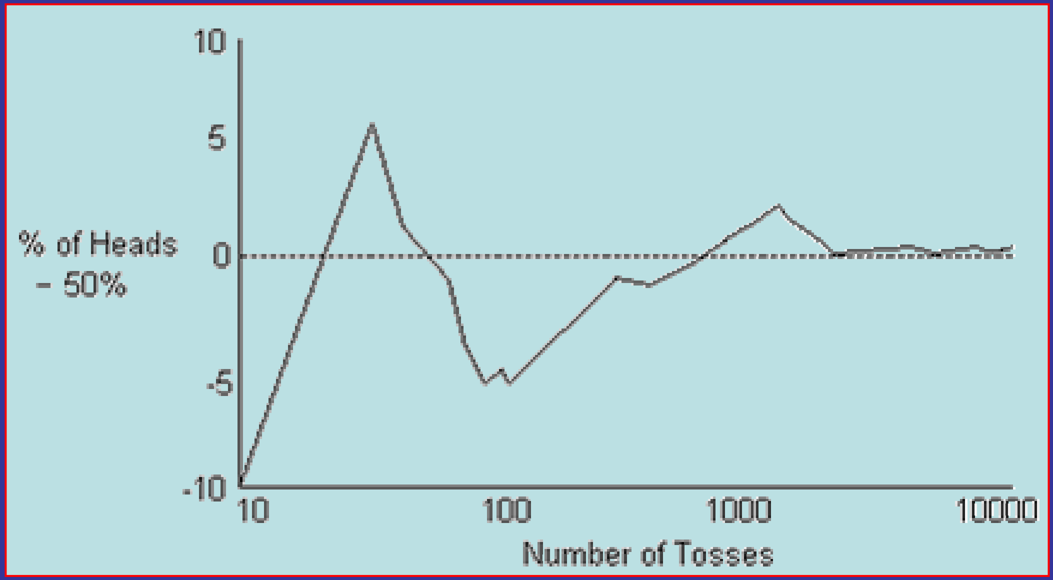


- A phenomenon is called *random* if the outcome of an experiment is uncertain.
- Random phenomena often follow recognizable patterns.
- This long-run regularity of random phenomena can be described mathematically.
- The mathematical study of randomness is called *probability theory*—probability provides a mathematical description of randomness.

10,000 Coin Tosses



- John Kerrich, a South African mathematician, was visiting Copenhagen when World War II broke out. Two days before he was scheduled to fly to England, the Germans invaded Denmark. Kerrich spent the rest of the war interned at a camp in Jutland and to pass the time he carried out a series of experiments in probability theory. In one, he tossed a coin 10,000 times. His results are shown in the following graph:



The horizontal axis is on a log scale

Introduction

- A tossed coin shows heads 50% of the time, on average. After many tosses, the number of heads (#heads) is approximately equal to the number of tails:

In the limit, as the number of tosses approaches ∞:

$$p(heads) = \frac{\#heads}{\#tosses} = \frac{1}{2}$$

- We say the probability of getting a head at any one toss is 1/2. This illustrates the concept of probability that will be used in this course.

Example—2 Coin Tosses

- Toss a coin twice and record for each toss whether the result was a head (H) or a tail (T). What are the possible outcomes?
  - HH, HT, TH, TT
- Let A be the event of one or more heads. Which outcomes belong to event A?
  - HH, HT, TH
- Let B be the event that there are no heads. Which outcomes belong to event B?
  - TT
- In this example, events A and B are said to be *disjoint* or *mutually exclusive*, as they have no outcomes in common. They are also *exhaustive*, as they cover all possible outcomes.
- Exercise: Define an event C which is not disjoint from A.
  - C is the event of exactly one head
  - C is the event of exactly two heads

Experiments and Events

- Simple Experiment:
  - *A simple experiment is some well-defined act or process that leads to a single well-defined outcome.*
- Typically, experiments have the following characteristics:
  - The set of all possible outcomes is known *before* the experiment.
  - The outcome of the experiment *is not known* beforehand.

Events as Sets of Possibilities

- The symbols *S* or *U* are often used to represent the *sample space*, or the set of all elementary events for a simple experiment.
- sample space = *S* or *U*
    - A fair coin is tossed and the outcome observed, $S=\{H,T\}$.
    - An individual is randomly sampled and their sex noted, $U=\{F, M\}$.

Events as Sets of Possibilities

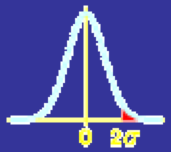
- Capital letters such as *A*, *B*, and so on represent events, each of which is a subset or sub-grouping of elementary events in *S*.
- Suppose we have two events, *A* and *B*, in *S*. Then we can consider an experimental outcome that is a member of both A and *B*—event *A* and *B* has occurred. This is symbolized as:
  - (A and B) = $(A \cap B)$
- Similarly, we can imagine an outcome that is *either* event *A* or *B*. This is known as the *union* of *A* and *B*, given as:
  - (A or B) = $(A \cup B)$
- By definition $(A \cap B)$ is also $(A \cup B)$, but, the reverse is not true.

Events as Sets of Possibilities

- Suppose that there is a simple experiment with sample space *S* and that *A* is some event in that sample space. Not only do we have a class of events *A*, but also *not* A. This is called the *complement*.

$$\text{Complement of } A = \bar{A}$$

Events as Sets of Possibilities



- In probability theory, any event that cannot be predicted with certainty, when measured, is called a *random variable*.
- A random variable is a *function* defined by the outcome of a random phenomenon.
  - If *S* is the sample space associated with some experiment, a random variable *X* is a function that assigns a real number *X(s)* to each element *s* that belongs to *S*.
- For random variables:
  - Each experiment or event has an outcome that cannot be determined before the event occurs.
  - However, each event or experiment must allow for the description of each and every possible outcome.

Events as Sets of Possibilities

- Suppose that an experiment consists of rolling two fair dice and observing the outcome.
- Let S be the sample space consisting of the 36 ordered pairs (i,j) where $1 \leq i \leq 6$ and $1 \leq j \leq 6$. Here, i is the integer resulting from the first die and j the integer resulting from the second die.
- We are interested in the sum of the two faces of the dice.
- The sum of the two faces of dice is a random variable.

Events as Sets of Possibilities

- Let X be the random variable indicating the sum of the two faces. Then, X can be represented by $X(i,j) = i+j$
- The *range space* of the random variable X is $\{2, 3, 4, \ldots, 12\}$ for a particular event.

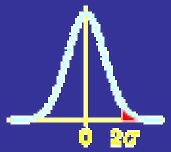
		DIE #1					
		1	2	3	4	5	6
D	1	2	3	4	5	6	7
I	2	3	4	5	6	7	8
E	3	4	5	6	7	8	9
	4	5	6	7	8	9	10
#	5	6	7	8	9	10	11
2	6	7	8	9	10	11	12

Events as Sets of Possibilities

- The collection of all outcomes is called the outcome or sample space or universe. Some examples of sample spaces are:
  - A rat from a cage is selected at random and its sex determined. The possible outcomes are male or female; the sample space is $S = \{M, F\}$.
  - Each of 6 students selects an integer between 1 and 52. We are interested in whether any two of the selected numbers match or are different; $S = \{M, D\}$.
  - A box of cereal contains 1 of 4 different prizes. Buying one box yields one prize as an outcome; $S = \{P1, P2, P3, P4\}$.
  - The state of Maryland institutes a daily lottery where a three-digit number is randomly selected as the winning number; $S = \{000, 001, 001, \ldots, 999\}$.
  - A fair coin is flipped until a head is observed. If K is equal to the number of trials necessary to produce a head, then K could be *any* positive integer; $S = \{1, 2, 3, \ldots \infty\}$
  - A biologist identifies a marsh bird. An adult bird is captured and weighed. $W =$ the weight of the bird. Based on past experience $S = \{200 \le W \le 450\}$.
  - Same biologist, same bird. Now however, the biologist determines the weight *and* sex of the bird, a two-dimensional sample space; $U = \{(S,W);\ S=M \text{ or } F,\ 200 \le W \le 450\}$.

Examples 1 through 4 show a finite number of outcomes.
Example 5, the number of outcomes is infinite but countable.
Example 6 shows outcomes represented by a range of possible outcomes.
Example 7 shows a two-dimensional sample space.

Elementary Probability Theory



- Probability is a set function P that assigns to each event *A* in the sample space *S* a number p(*A*), called the probability of event *A*, such that the following are true:
  - ❖ $p(A) \geq 0$ for every event *A*
  - ❖ $p(S) = 1.00$
  - ❖ If there is a set of countable events $\{A_1, A_2, \ldots, A_n\}$, and if the events are *mutually exclusive*, then:

$$p\left(A_1 \cup A_2 \cup \ldots \cup A_n\right) = p\left(A_1\right) + p\left(A_2\right) + \ldots p\left(A_n\right)$$

- The probability of the union of mutually exclusive events is the sum of their individual probabilities .

Probability Rules

The *rule of complementary probability*:

Rule 1a:

$$p(A) = 1 - p(\bar{A})$$

Given: $S = A \cup \bar{A}$ and $A \cap \bar{A} = \varnothing$

then: $1 = \mathrm{p}(\mathrm{A}) + \mathrm{p}(\bar{\mathrm{A}})$

therefore: $p(A) = 1 - p(\bar{A})$

Probability Rules

The *rule of complementary probability*:

Rule 1b:

$$p\left(\bar{A}\right) = 1 - p\left(A\right)$$

Given: $p\left(A \cup \bar{A}\right) = p\left(A\right) + p\left(\bar{A}\right)$

and: $p\left(A \cup \bar{A}\right) = S$ or 1

then: $p\left(A\right) + p\left(\bar{A}\right) = 1$

therefore: $p\left(\bar{A}\right) = 1 - p\left(A\right)$

Probability Rules

The *probability range*:

Rule 2:

For each event A, $P(A) \leq 1$

Since A is a subset of $S\,(A \subset S)$, then $P(A) \leq S$

Because $P(A) \geq 0$, then $0 \leq p(A) \leq 1$

Probability Rules

The *impossible event*:

Rule 3:

$$p(\varnothing) = 0, \text{ for any } S$$

The impossible event always has a probability of zero.

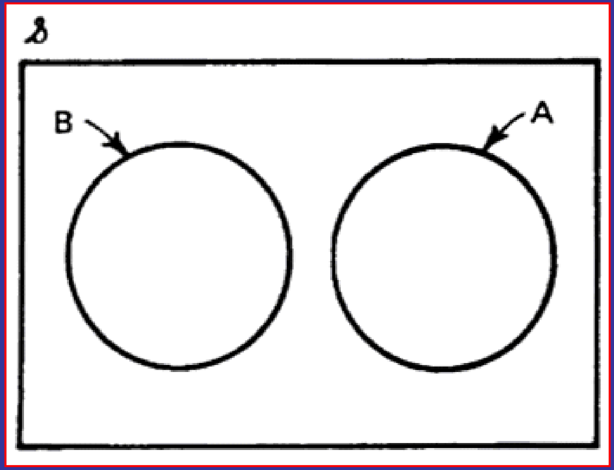
Probability Rules

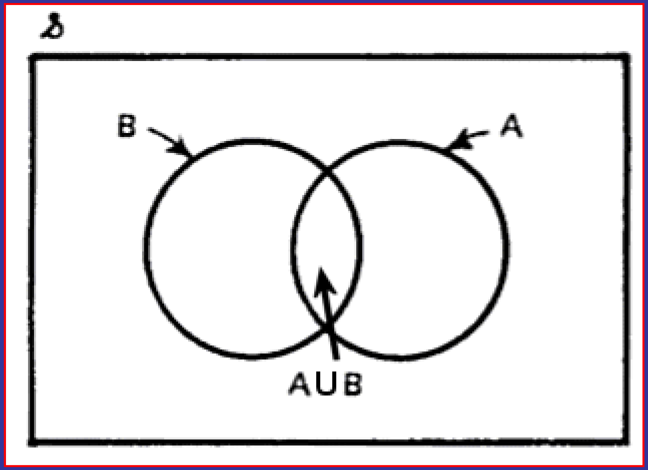
The *"or" rule of probability*:

Rule 4a:

For any two events, *A* and *B* in *S*:

$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$





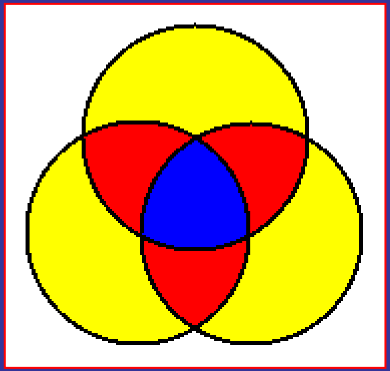
Probability Rules

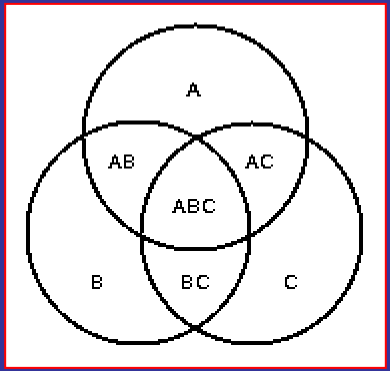
The *"or" rule of probability*:

Rule 4b:

For any three events, *A* and *B* in *S*:

$$p\left(A \cup B\right) = p\left(A\right) + p\left(B\right) + p\left(C\right) - p\left(A \cap B\right)$$
$$- p\left(A \cap C\right) - p\left(B \cap C\right) + p\left(A \cap B \cap C\right)$$





Probability Rules

The *"and" rule of probability*:

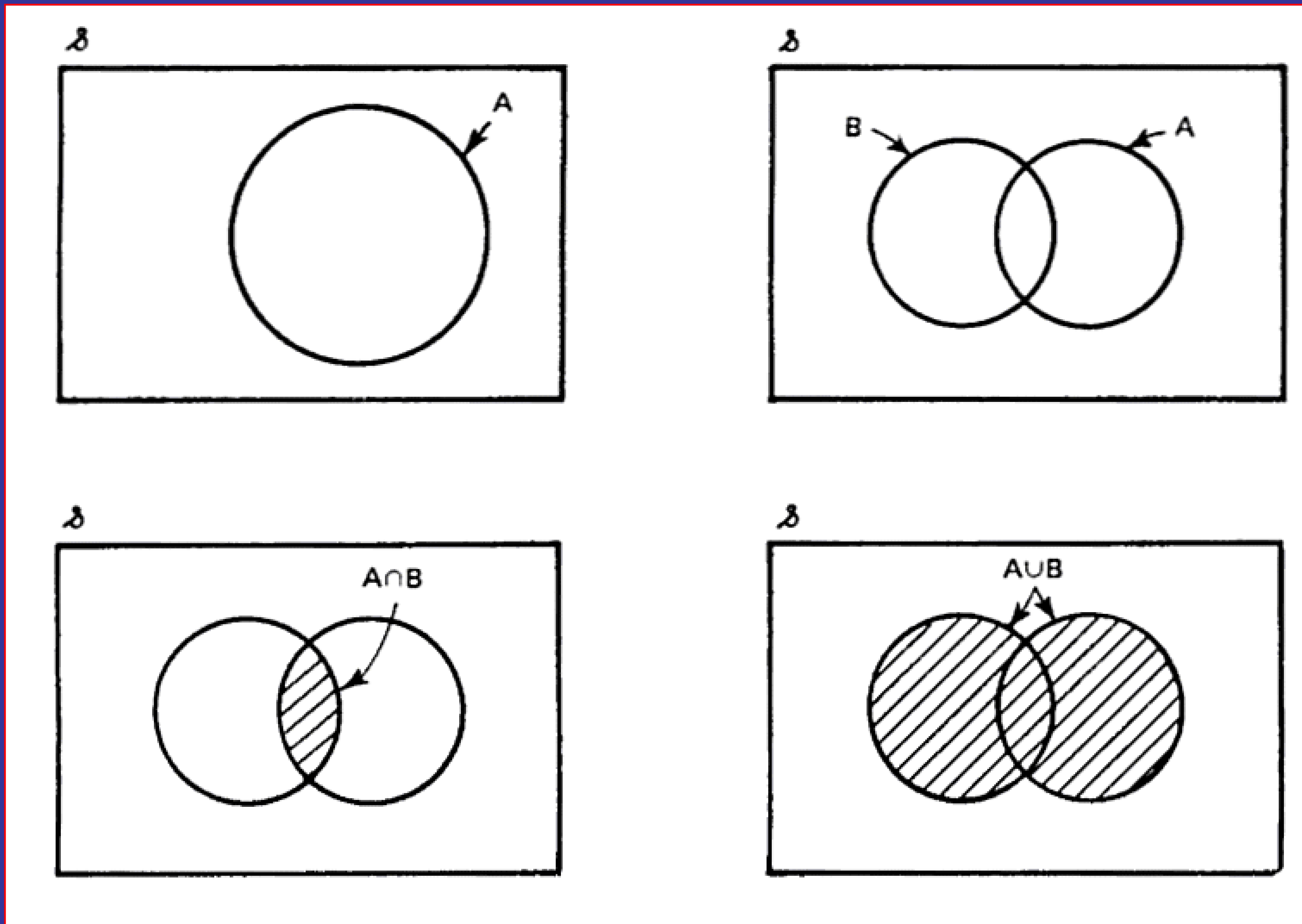
Rule 5:

When one trial has no influence on another, their joint probability is equal to:

$$p(A \cap B) = p(A)\,p(B)$$

$$P(H \cap H) = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{4}$$

Venn Diagrams



Venn Diagrams