Ch 9. Discrete Probability (1)

Sungwon Kang

Acknowledgement

- Michael P. Frank, University of Florida, COT 3100, Applications of Discrete Structures, Lecture Slides
- [Rosen 19] Kenneth H. Rosen, for Discrete Mathematics & Its Applications (8th Edition), Lecture slides
- [Hunter 11] David J. Hunter, Essentials of Discrete Mathematics, 2nd Edition, Jones & Bartlett Publishers, 2011, Lecture Slides
- Milos Hauskrecht, Pittsburgh University, Lecture Slides

Logically valid but not useful inference

Statement 1. If a transportation is unsafe,
then we should not use that transportation.
Statement 2. Any transportation that fails is unsafe.
Statement 3. Airplane fails.

Statement 4. We should not take an airplane.

Two Kinds of Inference (1/2)

Deductive Inference

Socrates is a human.

All humans die.

Socrates dies.

If a deductive inference is logically valid, then its conclusion must be accepted once the premises are accepted.

If leaded gasoline is safe, then its use for cars should be permitted.

Leaded gasoline is safe.

The use of leaded gasoline for cars should be permitted.

Two Kinds of Inference (2/2)

Deductive Inference

Socrates is a human.
All humans die.

Socrates dies.

If a deductive inference is logically valid, then its conclusion must be accepted once the premises are accepted.

If leaded gasoline is safe, then its use for cars should be permitted.
Leaded gasoline is safe.

The use of leaded gasoline for cars should be permitted.

Inductive Inference

That swan is white.
Another swan is white, too.
This is a swan.

This swan must be white.

The conclusion is not guaranteed but can be useful and sometime “rational”.

Why Probability?

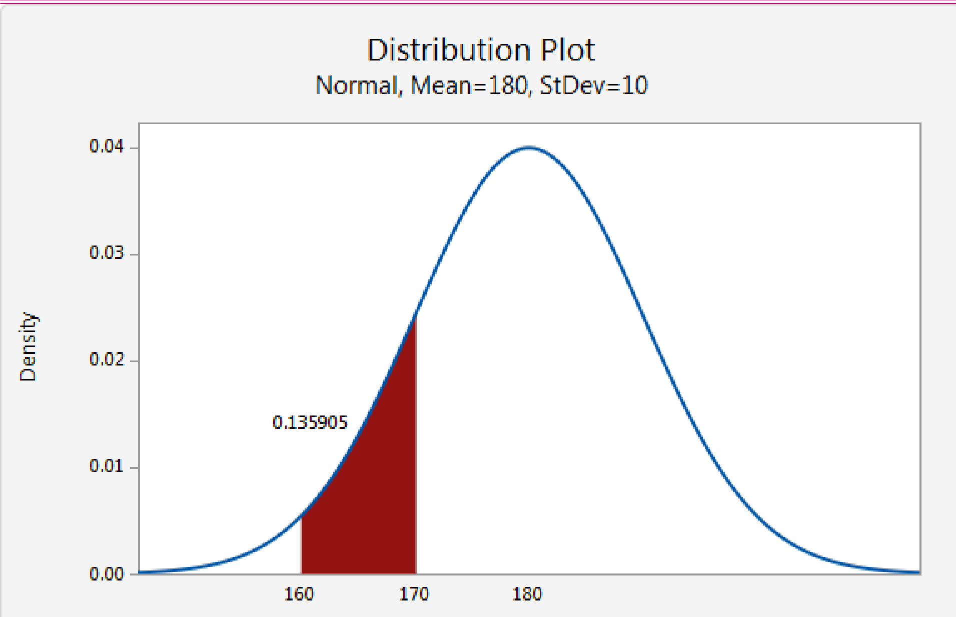
- In the real world, we often do not know whether a given proposition is true or false.

- Probability theory provides us a way to reason about propositions, of which the truths are *uncertain*.

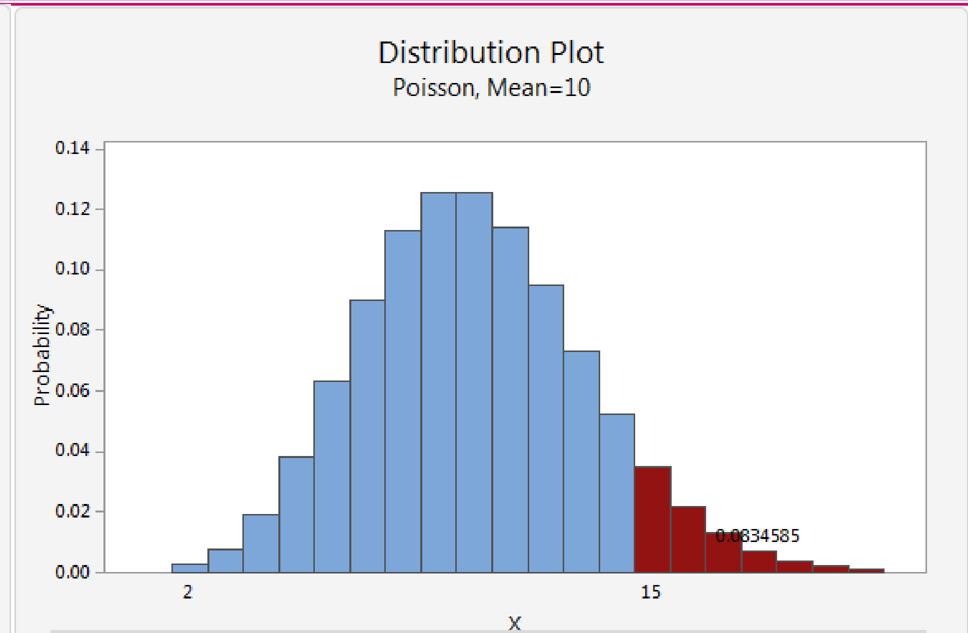
- Useful in
  - Weather forecasting: 80% Probability of precipitation (chance of rain)
  - Baseball: Batting Average of Babe Ruth 0.342
  - Gambling (> 300 years of application history)
  - weighing evidence
  - diagnosing problems

- In general, useful for analyzing situations whose exact details are unknown

Continuous vs. discrete probability



X: continuous random variable
represents continuous outcomes



☛ **X:** discrete random variable
represents discrete outcomes

- A discrete random variable is a random variable that is **countable**, such as a list of non-negative integers.

  ☛ You should know how to count in order to calculate discrete probability.
- A discrete distribution describes the probability of each value of a discrete random variable.

Ch 7. Discrete Probability

Table of Contents

1. Finite probability space, events

Experiment: a procedure that yields one of the possible outcomes

(Finite) Sample space: a (finite) set of all possible outcomes

Event**:** a subset of possible outcomes

(Event E is a subset of a sample space S)

Note) “Outcome” is an ”actualization” of an event or

a concrete event for an (abstract) event.

Example (Experiment of rolling two dice)

What is the probability of an event that

(the sum of the values of) the outcome is 7?

All possible outcomes (sample space S):

(1,1) (1,2)... (1,6) (2,1) (2,2) … (2, 6) ... (6,1) (6,2) … (6,6)

Outcomes leading to 7 (event E)

(1,6) (2,5) …(6,1)

Probability: definitions

Definition (by Laplace, 1749-1827)

Suppose A is a subset of a nonempty finite set U. The *probability* that a randomly chosen element of U lies in A is the ratio

$$P(A) = \frac{|A|}{|U|}.$$

The set U is called the *sample space*, and the set A is called an *event*.

It is assumed that the outcomes are equally likely unless we say otherwise.

Example (Experiment of rolling two dice)
What is the probability of an event that
(the sum of the values of) the outcome is 7?

All possible outcomes (sample space S):

(1,6) (2,6) …(6,1),…(6,6) => count: 36

Outcomes leading to 7 (event E)

(1,6) (2,5) …(6,1) => count: 6

P(sum=7)= 6/36 = 1/6

Example Toss a coin 10 times, and it lands on Heads each time.
(Heads and tails are equality likely by assumption.)
What is the probability of this event?

HHHHHHHHHH

Example Toss a coin 10 times, and it lands on Heads each time.
(Heads and tails are equality likely by assumption.)
What is the chance that another heads will come up on the next toss?

HHHHHHHHHH ___
?

< 0.5 ?

= 0.5 ?

> 0.5 ?

Probability: simple examples

Example: Suppose you get a random license plate from all the possible Illinois plates described above. What is the probability that your plate contains the word `CUB` or the word `SOX`?

Illinois plate consists of 3 letters followed by 3 digits or
2 letters followed by 4 digits.

What is $|U|$?

What is $|A|$?

Example: Suppose you get a random license plate from all the possible Illinois plates described above. What is the probability that your plate contains the word `CUB` or the word `SOX`?

Solution

From above, $|U| = 24,336,000.$. *Let A be the event that a plate contains the words CUB or SOX.*

$$P(A) = \frac{|A|}{|U|} = \frac{10^3 + 10^3}{24,336,000} = \frac{1}{12168} \approx 0.000082$$

[redacted]

Example: If you roll two standard six-sided dice, what is the probability that you roll an 8 (i.e., that the sum of the values on the two dice will be 8)?

What is |U|?

What is |E|?

[redacted]

Example: If you roll two standard six-sided dice, what is the probability that you roll an 8 (i.e., that the sum of the values on the two dice will be 8)?

Solution

The size of the sample space is $|D_1 \times D_2| = |D_1| \cdot |D_2| = 36$*. The following ordered pairs*

$$\{(2,6),(3,5),(4,4),(5,3),(6,2)\}$$

represent the event of rolling an 8. Hence the probability of such a roll is $5/36 \approx 0.139$*.*

2. Basic Concepts of Probability Theory

1) Probability of an Event
2) Axioms of probability, Probability measures
3) Probability of a Complement of an Event
4) Probability of a Union of Events
5) Probability Distribution

1) Probability of an Event

Definition:

The probability of an event E is

the sum of the probabilities of the outcomes in E.

$$P(E) = \sum_{s \in E} P(s)$$

Example:

What is the probability that an odd number appears when we roll a die?

Solution: We want the probability of the event E = {1,3,5}.

Probabilities of outcomes:

P(1) = P(3) = P(5) = 1/6.

Then, $P(E) = \sum_{s \in E} P(s)$

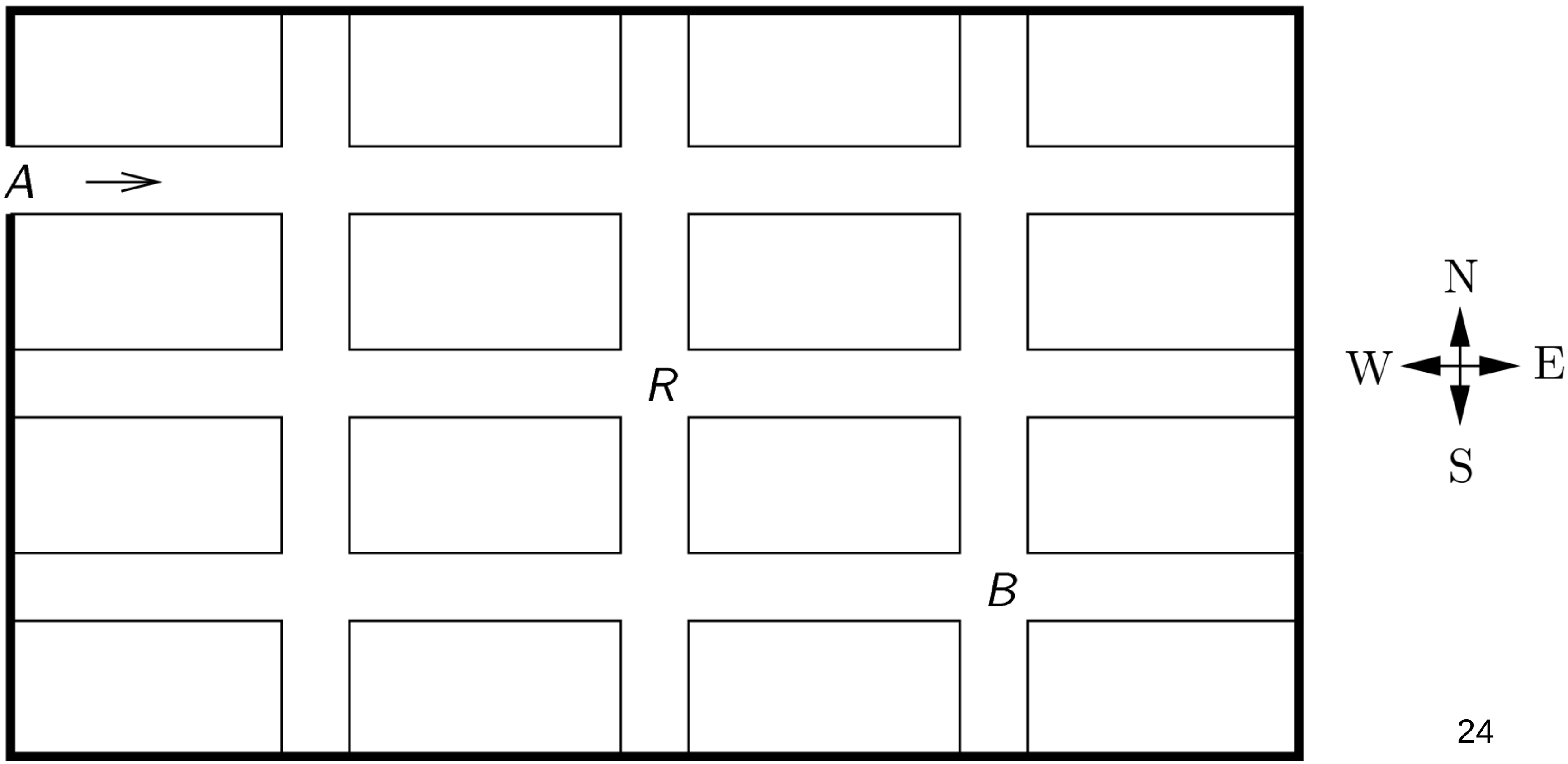
$= P(1) + P(3) + P(5)$

$= 1/6 + 1/6 + 1/6$

$= 3/6$

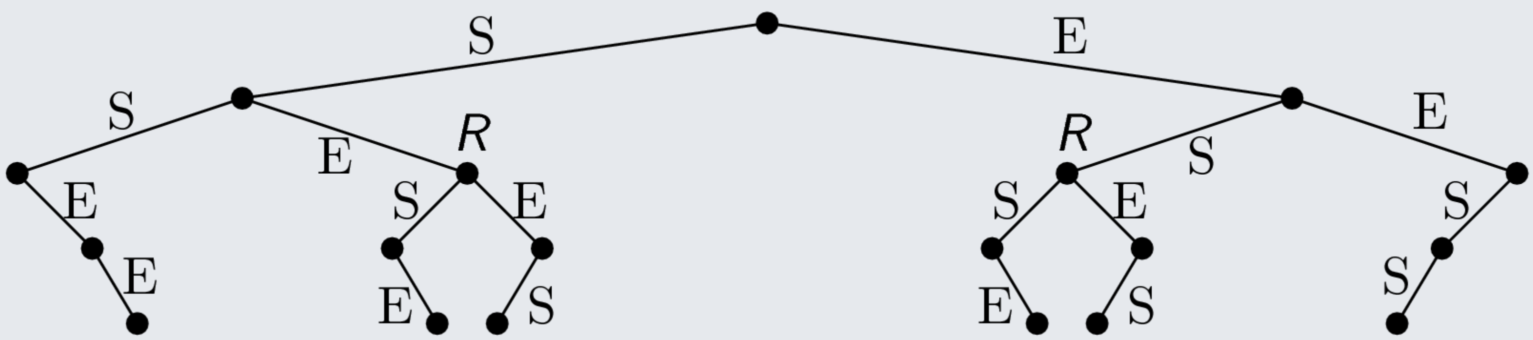
Application: locating a restaurant

Starting at A, what is the probability that you pass restaurant R on the way (directly) to point B?



Solution

Use the decision tree:



Of the six possible direct paths, four pass through R, so the desired probability is $4/6 = 2/3$.

Application: sampling

Rodelio wants to know if a majority of the voters in his town will support his candidacy for mayor. Suppose, for the sake of this discussion, that out of the 300 voters in this town, 151 support Rodelio (but Rodelio doesn't know this information). Rodelio selects 20 voters at random from the population of 300. What is the probability that, out of this random sample, fewer than five support Rodelio?

Solution

The sample space U *is the set of all possible random samples, so* $|U| = C(300, 20)$. *The event* A *that fewer than five of the voters in this sample support Rodelio is* $|A|$

$$\begin{aligned} = \quad & C(149,20) + C(151,1)\cdot C(149,19) + C(151,2)\cdot C(149,18) \\ & + C(151,3)\cdot C(149,17) + C(151,4)\cdot C(149,16) \end{aligned}$$

The desired probability is $P(A) = |A|/|U| \approx 0.0042$.

Now suppose that Rodelio conducts a poll of a sample of 20 voters, and finds that only four supporters support his candidacy. Do you like his chances to win the election?

2) Axioms of probability, Probability measures

Three axioms of the Probability Theory

Axiom 1. The probability of a discrete outcome s satisfies:

$$0 \le P(s) \le 1$$

Axiom 2. The sum of probabilities of all (disjoint) outcomes is 1.

Axiom 3. For any two events E1 and E2,

$$P(E1 \cup E2) = P(E1) + P(E2) – P(E1 \cap E2)$$

☛ Probability Version of the **Inclusion-Exclusion Principle**

$$|E1 \cup E2| = |E1| + |E2| – |E1 \cap E2|$$

Probability measures

- A probability measure is a real-valued function defined on a set of events in a probability space.

Examples

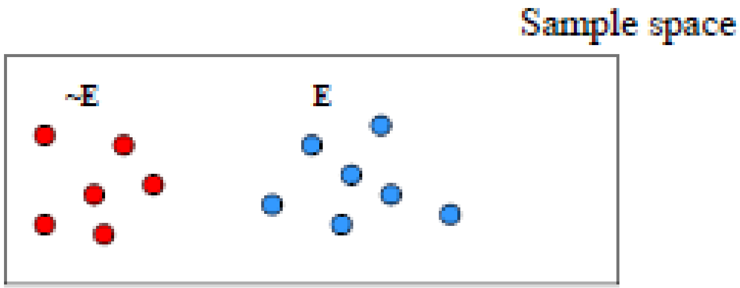
- complement
- union
- probability distribution
- etc.

See the following subsections.

3) Probability of a Complement of an Event

Theorem Let E be an event and ~E its complement with regard to S. Then

$$P(\sim E) = 1 - P(E)$$



Proof

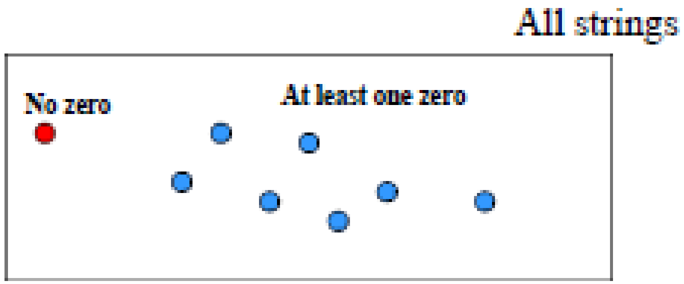
$P(\sim E) = (|S| - |E|) / |S|$

$= 1 - |E| / |S|$

Example

Consider a string with 10 randomly generated bits.

What is the probability that there is at least one zero in the string?



Event: seeing no-zero string P(E) =?

~Event: seeing at least one zero in the string ?

Event: seeing no-zero string P(E) = $1/ 2^{10}$

~Event: seeing at least one zero in the string

P(~E) = 1-P(E) = $1- 1/2^{10}$

Example: If you roll two standard six-sided dice, what is the probability you roll 10 or less?

Example: If you roll two standard six-sided dice, what is the probability you roll 10 or less?

Solution

Compute the probability of rolling more than 10: There are two ways to roll 11 and one way to roll 12, so the probability of rolling more than 10 is $(2+1)/36 = 1/12$*. Thus the probability of rolling 10 or less is* $1 - 1/12 = 11/12$*.*

Application: quality control

Suppose that there are 10 defective machines in a group of 200. A quality control inspector takes a sample of 3 machines and tests them for defects. How likely is it that the inspector discovers a defective machine?

Suppose that there are 10 defective machines in a group of 200. A quality control inspector takes a sample of 3 machines and tests them for defects. How likely is it that the inspector discovers a defective machine?

Solution

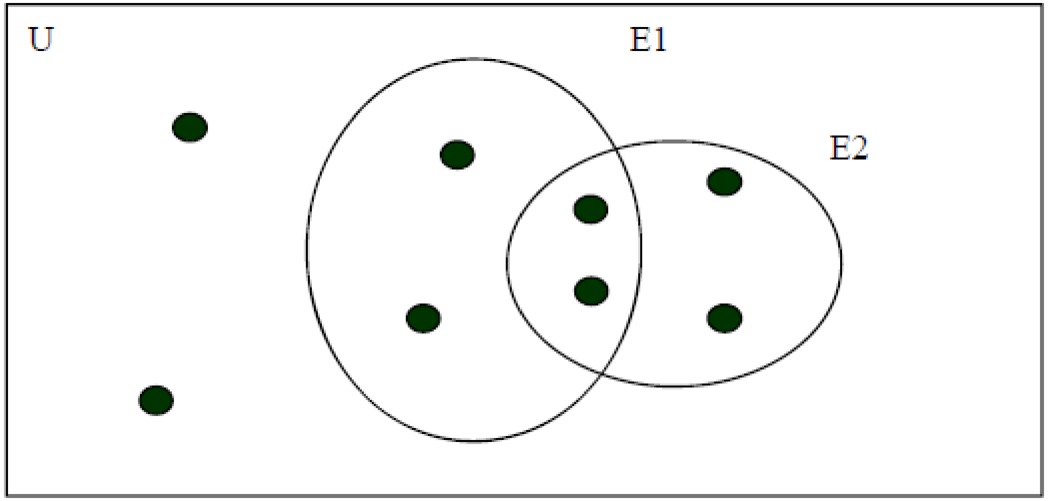
The sample space is the total number of selections of three machines: $C(200, 3)$. *The event that at least one of the machines is defective is the opposite of the event that none are. There are 190 non-defective machines, so* $C(190, 3)$ *samples contain no defects. Therefore the desired probability is*

$$1 - \frac{C(190, 3)}{C(200, 3)} = 1 - \frac{1,125,180}{1,313,400} \approx 0.1433.$$

4) Probability of a Union of Events

Theorem Let E1 and E2 be two events in the sample space S. Then

$$P(E1 \cup E2) = P(E1) + P(E2) - P(E1 \cap E2)$$



Example What is the probability that a positive integer <= 100 is divisible either by 2 or 5?

(Let E1 be the event of "an integer being divisible by 2" and

E2 be the event of "an integer being divisible by 5".)

$P(E1) = 50/100$

$P(E2) = 20/100$

$P(E1 \cap E2) = 10/100$

$P(E1 \cup E2) = 50/100 + 20/100 - 10/100$

$= 60/100$

Disjoint (Mutually Exclusive) & Exhaustive Set of Events

- Two events E_1 and E_2 are disjoint (or mutually exclusive) if $E_1 \cap E_2 = \varnothing$
  - Two mutually exclusive events **cannot both occur** in the same instance of a given experiment.
  - For disjoint events E1 and E2,
$$P(E_1 \cup E_2) = P(E_1) + P(E_2).$$

- A set $E = \{E_1, E_2, \ldots\}$ of events in the sample space S is *exhaustive* if
$$\bigcup E_i = S$$

- An exhaustive set of disjoint events has the property that
$$\sum P(E_i) = 1$$

Disjoint vs. Complementary

- Do the sum of probabilities of two disjoint events always add up to 1?

- Do the sum of probabilities of two complementary events always add up to 1?

Disjoint vs. Complementary

- Do the sum of probabilities of two disjoint events always add up to 1?

  – Not necessarily.
  There may be more than 2 events in the sample space.

- Do the sum of probabilities of two complementary events always add up to 1?

  - Yes. By the definition of “complementary”.

5) Probability Distribution

- So far we assumed that
  the probabilities of all outcomes are equally likely.
  - ☛ Uniform Distribution Assumption
- However, in many cases outcomes may not be equally likely.

Example 1

Biased Coin:

– Probability of heads: 0.6

– Probability of tails: 0.4

Example 2

Biased Die:

– Probability of 6: 0.4

– Probability of 1, 2, 3, 4, 5: 0.12 each

Definition: A function p: S $\rightarrow$ [0,1] is *a probability distribution* if :

1) The events of S are disjoint.
2) The probability of each event is $\geq$0 and $\leq$1.
3) The sum of the probabilities is 1.

Example See the previous slide.

Quiz 20-1

Suppose you draw five cards from a deck of cards that you have seen when we study counting. What is the probability of getting three cards (but not four cards) of one kind?

(a) C(13,1) C(4,1) C(48,1) / C(52,5)
(b) C(13,1) C(4,1) C(48,2) / C(52,5)
(c) C(13,1) C(4,4) C(48,1) / C(52,5)
(d) C(13,1) C(4,4) C(48,2) / C(52,5)
(e) C(13,3) C(4,1) C(48,1) / C(52,5)
(f) C(13,3) C(4,1) C(48,2) / C(52,5)
(g) C(13,3) C(4,4) C(48,1) / C(52,5)
(h) C(13,3) C(4,4) C(48,2) / C(52,5)