Chapter 2
Probability

吳明龍 副教授
成大資訊系/醫資所
minglong.wu@csie.ncku.edu.tw
辦公室:資訊系館 12 樓

2.1 Sample Space

- In the study of statistics, we are concerned with the presentation and interpretation of chance outcomes.
- Observation: Any recording of information, whether it be numerical or categorical, is referred to observation.
  - 2, 0, 1, 2
  - D, N, D, N, N
- 'Experiment' is used to describe any process that generates a set of data.
  - E.g., tossing of a coin, two possible outcomes, heads and tails
- In most cases the outcome will depend on chance and, thus, cannot be predicted with certainty.
- Definition 2.1: The set of possible outcomes of a statistical experiment is called the **sample space**, represented by **S**.
- Each outcome in a sample space is called an element, a member, or a sample point.

Sample Space

- Example 2.1
  - Tossing a coin: $S = \{H, T\}$
  - Tossing a die
    - $S_1 = \{1, 2, 3, 4, 5, 6\}$
    - $S_2 = \{even, odd\}$
- Tree diagram: List the elements of the sample space systematically.
- Example 2.2
  - $S = \{HH, HT, T1, T2, T3, T4, T5, T6\}$
- Example 2.3
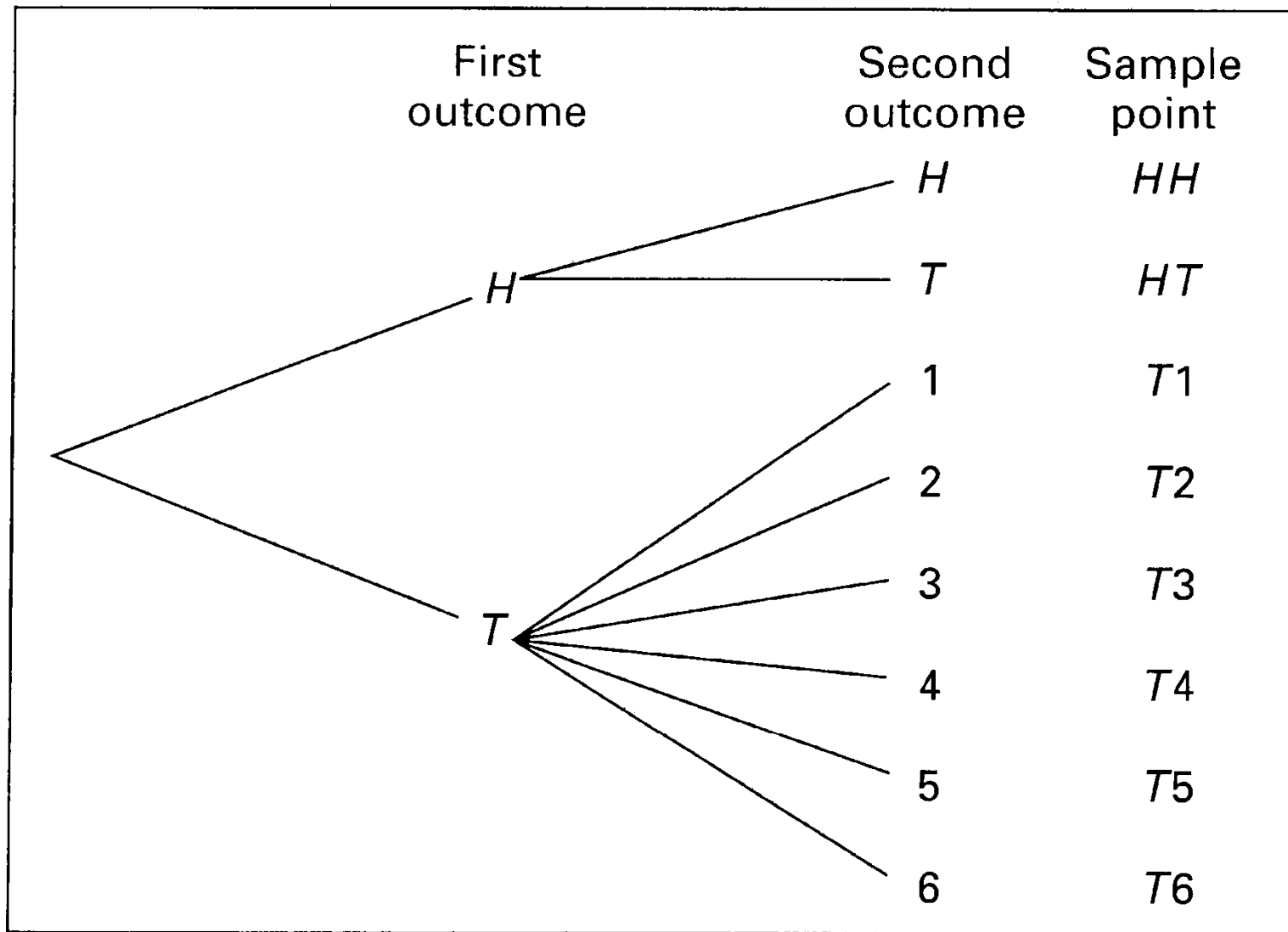 - $S = \{DDD, DDN, DND, DNN, NDD, NDN, NND, NNN\}$

Figure 2.1 Tree diagram for Example 2.2.

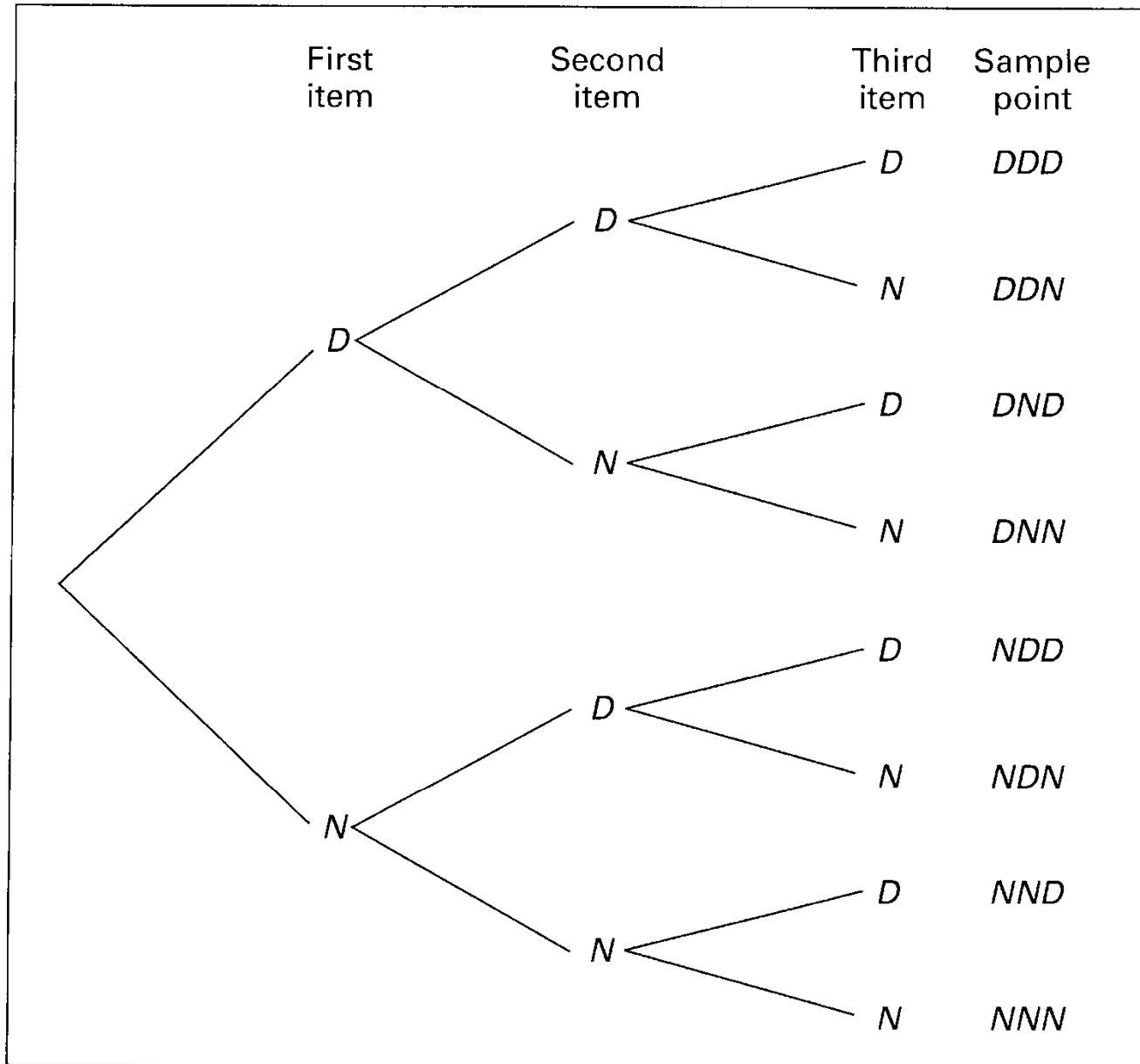**Figure 2.2** Tree diagram for Example 2.3.

Sample Space

- Statement (Rule): Sample spaces with a large or infinite number of sample points are best described by a statement or rule.
  - $S = \{x | x$ is a city with population over 1 million$\}$
  - $S = \{(x,y) | x^2 + y^2 \leq 4\}$
- The rule method has practical advantages, particularly for the many experiments where a listing becomes a tedious chore (雜物).

2.2 Events

- Definition 2.2: An <u>event</u> is a subset of a sample space.
  - E.g., we may be interested in the event A that the outcome when a die is tossed is divisible by 3, $A = \{3, 6\}$.
- Example 2.4
  - Given the sample space $S = \{t \mid t \geq 0\}$, where t is the life in years of a certain electronic component.
  - The event A that the component fails before the end of the fifth year is the subset $A = \{t \mid 0 \leq t < 5\}$.
- <u>Null set</u>, denoted Ø, contains no elements at all.

Events

- Definition 2.3: The <u>complement</u> of an event A with respect to S is the subset of all elements of S that are not in A. We denote the complement of A by the symbol <u>A'</u>.

- Example 2.5
  - Let R be the event that a red card is selected from an ordinary deck of 52 playing cards.
  - S be the entire deck.
  - R' is the event that the card selected from the deck is not a red but a black card.

Events

- Definition 2.4: The intersection of two events A and B, denoted by the symbol $A \cap B$, is the event containing all elements that are common to A and B.

- Example 2.7
  - Let E be the event that a person selected at random in a classroom is majoring in engineering.
  - F is the event that the person is female.
  - The event $E \cap F$ is the set of all female engineering students in the classroom.

Events

- Definition 2.5: Two events *A* and *B* are mutually exclusive, or disjoint if $A \cap B = \varnothing$, i.e., if *A* and *B* have no elements in common.

- Definition 2.6: The union of the two events *A* and *B*, denoted by the symbol $A \cup B$, is the event containing all the elements that belong to *A* or *B* or both.

Events

- Venn diagram: The relationship between <u>events</u> and the corresponding <u>sample space</u> can be illustrated graphically by Venn diagram.

- In a Venn diagram, let the sample space be a <u>rectangle</u> and represent events by <u>circles</u>.

- In Figure 2.3

  – $A \cap B$ = regions 1 and 2

  – $A \cup C$ = regions 1, 2, 3, 4, 5, and 7
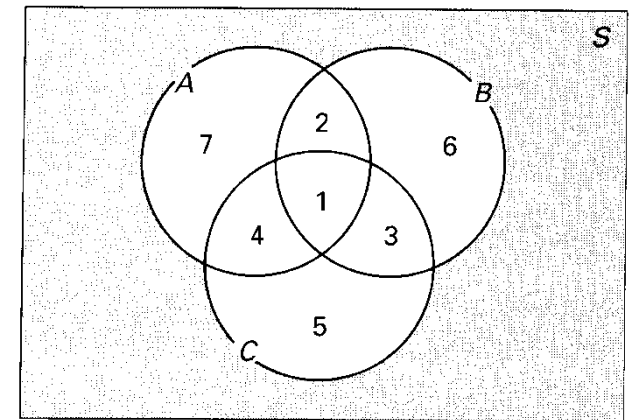
  – $B' \cap A$ = regions 4 and 7

Figure 2.3 Events represented by various regions.

2.3 Counting Sample Points

- Multiplication rule (Rule 2.1): If an operation can be performed in n_1 ways, and if for each of these a second operation can be performed in n_2 ways, then the two operations can be performed together in n_1n_2 ways.

- Example 2.13
  - Sampling points of a dice thrown twice? $n_1n_2 = (6)(6) = 36$ possible ways.

Counting Sample Points

- Example 2.14
    - Offers for home buyers with Tudor, rustic, colonial, and traditional exterior styling
    - With ranch, two-story, and split-level floor plans.
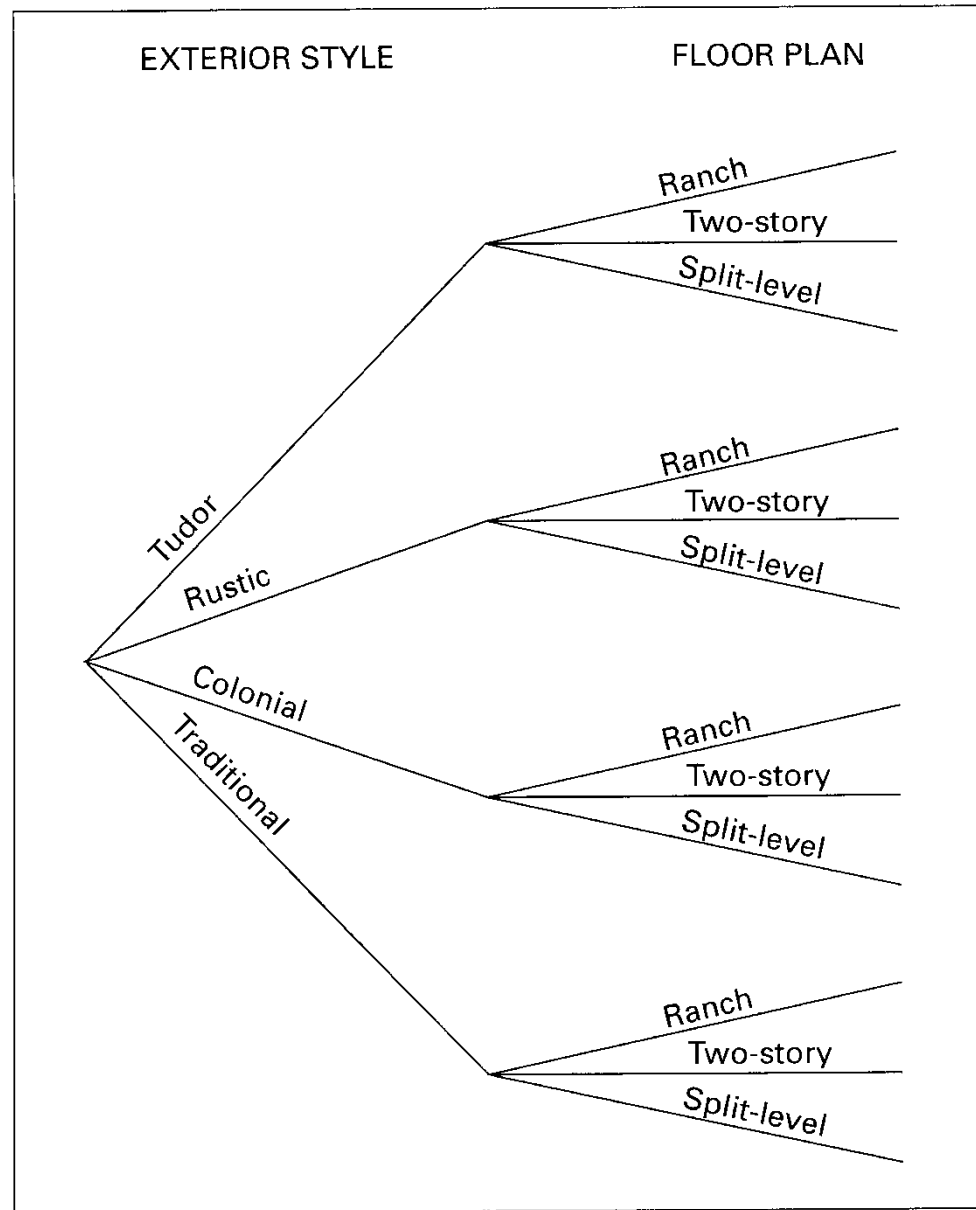   - How many ways can a buyer order one of these homes?

Figure 2.6 Tree diagram for Example 2.14.

Counting Sample Points

- Rule 2.2:
  - If an operation can be performed in n_1 ways, and if for each of these a second operation can be performed in n_2 ways, and for each of the first two a third operation can be performed in n_3 ways, and so forth, then the sequence of k operations can be performed in $n_1 n_2 \ldots n_k$ ways.

Counting Sample Points

- Example 2.16
  - Sam is going to assemble a computer by himself. He has the choice of chips from two brands, a hard drive from four, memory from three, and an accessory bundle from five local stores. How many different ways can Sam order the parts?
  - n_1 x n_2 x n_3 x n_4 = 2 x 4 x 3 x 5 = 120

Counting Sample Points

- Example 2.17
  - How many even four-digit numbers can be formed from the digits 0, 1, 2, 5, 6 and 9 if each digit can be used only once?
  - Unit digit 0
    - $n_1 n_2 n_3 n_4 = (1)(5)(4)(3) = 60$
  - Unit digit 2, 6
    - $n_1 n_2 n_3 n_4 = (2)(4)(4)(3) = 96$
  - Total numbers = 60 + 96 = 156

Counting Sample Points

- A <u>permutation</u> is an arrangement of all or part of a set of objects.

- A <u>combination</u> is the number of ways of selecting *r* objects from *n* without regard to order.

- Definition 2.8

  – For any non-negative integer n, n!, called "n factorial," is defined as n! = n(n-1)…(2)(1),

  with special case 0! = 1.

Counting Sample Points

- The number of permutations of n distinct objects is $n!$.
- The number of permutations of n distinct objects taken r at a time is
$$P(n,r) = n \times (n-1) \times ... \times (n-r+1) = \frac{n!}{(n-r)!}.$$
- The number of permutations of n distinct objects arranged in a circle is $\frac{n!}{n} = (n-1)!.$
- The number of permutations of n things of which n_1 are of one kind, n_2 of a second kind, …, and n_k of a kth kind is $\frac{n!}{n_1!n_2!...n_k!}.$
- The number of arrangements of a set of n objects into r cells with n_1 elements in the first cell, n_2 elements in the second, and so forth, is

$$\binom{n}{n_1,n_2,\cdots,n_r} = \frac{n!}{n_1!n_2!...n_r!}, \text{ where } n_1 + n_2 + \cdots + n_r = n.$$
- The number of combinations of n distinct objects taken r at a time is
$$C(n,r) = \binom{n}{r} = \frac{n!}{r!(n-r)!}.$$

Counting Sample Points

- ## Example 2.21

  - In how many ways can 7 graduate students be assigned to 1 triple and 2 double hotel rooms during a conference?

  -

    $$\binom{7}{3,2,2} = \frac{7!}{3!2!2!} = 210$$

- ## Example 2.23

  - How many difference letter arrangements can be made from the letters in the word STATISTICS?

    $$\binom{10}{3,3,2,1,1} = \frac{10!}{3!3!2!1!1!} = 50400$$

2.4 Probability of Event

- Perhaps it was <u>man's unquenchable (不能遏制的) thirst for gambling</u> that led to the early development of probability theory.
- What do we mean when we make the statements
  - John will <u>probably</u> win the tennis match.
  - I have a <u>fifty-fifty chance</u> of getting an even number when a die is tossed.
  - I am not <u>likely</u> to win at bingo tonight.
  - Most of our graduating class will <u>likely</u> be married within 3 years.
- In each case, we are expressing an outcome of which we are <u>not certain</u>, but owing to <u>past information</u> or from an understanding of the <u>structure of the experiment</u>, we have some <u>degree of confidence</u> in the <u>validity of the statement</u>.

Probability of Event

- The likelihood of the occurrence of an event resulting from a statistical experiment is evaluated by means of a set of real numbers called <u>weights</u> or <u>probabilities</u>, and ranged from 0 to 1.

- To every point in the sample space we assign a probability such that the sum of all probabilities is 1.

- In many experiments, such as tossing a coin or a die, all the sample points have the same chance of occurring and are assigned <u>equal probabilities</u>.

- For points outside the sample space, i.e., for simple events that cannot possibly occur, we assign a <u>probability of zero</u>.

Probability of Event

- Definition 2.9: The probability of an event A is the sum of the weights of all sample points in A.

$$0 \leq P(A) \leq 1, \quad P(\phi) = 0, \quad \text{and} \quad P(S) = 1$$

- If A_1, A_2, A_3,...is a sequence of <u>mutually exclusive events</u>, then

$$P(A_1 \cup A_2 \cup A_3 \cup \cdots) = P(A_1) + P(A_2) + P(A_3) \cdots$$

- Example 2.24: A coin is tossed twice. What is the probability that at least one head occur?

$$S = \{HH, HT, TH, TT\}, A = \{HH, HT, TH\}, \text{ and } P(A) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}$$

Probability of Event

- If an experiment can result in any one of N different equally likely outcomes, and if exactly n of these outcomes correspond to event A, then the <u>probability of event A</u> is $P(A) = \dfrac{n}{N}$.

- Example 2.28: In a poker hand consisting of 5 cards, find the probability of holding 2 aces and 3 jacks.

$$P(C) = \frac{C(4,2) \times C(4,3)}{C(52,5)} = \frac{\dfrac{4!}{2!2!} \times \dfrac{4!}{3!1!}}{\dfrac{52!}{5!47!}} = \frac{24}{2,598,960} \approx 0.9 \times 10^{-5}$$

Probability of Event

- If the outcomes of an experiment are not equally likely to occur, the probabilities must be assigned based on prior knowledge or experimental evidence.

- According to the relative frequency definition of probability, the true probabilities would be the fractions of events that occur in the long run.

- The use of intuition, personal beliefs, and other indirect information in arriving at probabilities is referred to as the subjective definition of probability.

2.5 Additive Rules

- Theorem 2.7: If *A* and *B*
  are any two events,
  then $P(A \cup B) = P(A) + P(B) - P(A \cap B).$

- Corollary 2.1: If *A* and *B*
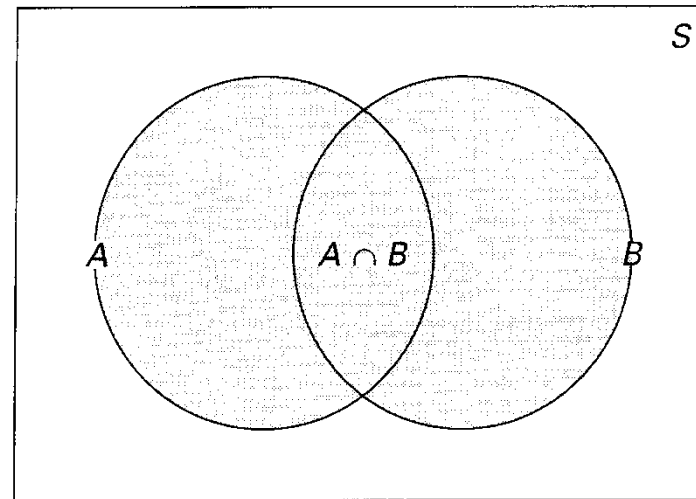 are <u>mutually exclusive</u>,
  then

  $P(A \cup B) = P(A) + P(B).$

Figure 2.7 Additive rule of probability.

Additive Rules

- Corollary 2.2: If A_1, A_2, …, A_n, are <u>mutually exclusive</u>, then

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n).$$

- Corollary 2.3: If A_1, A_2, …, A_n, is a partition of a sample space S, then

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n)$$

$$= P(S)$$

$$= 1.$$

- Theorem 2.8: For three events A, B, and C,

$$P(A \cup B \cup C) = P(A) + P(B) + P(C)$$

$$- P(A \cap B) - P(A \cap C) - P(B \cap C)$$

$$+ P(A \cap B \cap C).$$

Additive Rules

- Theorem 2.9: If *A* and *A'* are complementary events, then *P*(*A*)+*P*(*A'*)=1
  - Proof

$$\because A \cup A' = S, \text{ and } A \text{ and } A' \text{ are disjoint}$$

$$\therefore 1 = P(S) = P(A \cup A') = P(A) + P(A')$$

Additive Rules

- Example 2.32
  - If the probabilities that an automobile mechanic will service 3, 4, 5, 6, 7, or 8 or more cars on any given workday are, respectively 0.12, 0.19, 0.28, 0.24, 0.10, and 0.07, what is the probability that he will service at least 5 cars on his next day at work?
  - E: at least 5 cars serviced.
  - $P(E) = 1 - P(E') = 1 - (0.12 + 0.19) = 0.69$