Chapter 2: Probability

Department of Engineering Sciences
Izmir Katip Celebi University

Week 3
2014-2015 Spring

Sample Space and Events

In the study of statistics, we are concerned basically with the presentation and interpretation of **chance outcomes** that occur in a planned study or scientific investigation. For example, we may record the number of accidents that occur monthly at the intersection of Driftwood Lane and Royal Oak Drive, hoping to justify the installation of a traffic light; we might classify items coming off an assembly line as "defective" or "nondefective"; or we may be interested in the volume of gas released in a chemical reaction when the concentration of an acid is varied. Hence, the statistician is often dealing with either **numerical data**, representing counts or measurements, or **categorical data**, which can be classified according to some criterion.

We shall refer to any recording of information, whether it be numerical or categorical, as an **observation**. Thus, the numbers 2, 0, 1, and 2, representing the number of accidents that occurred for each month from January through April during the past year at the intersection of Driftwood Lane and Royal Oak Drive, constitute a set of observations. Similarly, the categorical data N, D, N, N, and D, representing the items found to be defective or nondefective when five items are inspected, are recorded as observations.

Statisticians use the word experiment to describe any process that generates a set of data. A simple example of a statistical experiment is the tossing of a coin. In this experiment, there are only two possible outcomes, heads or tails. Another experiment might be the launching of a missile and observing of its velocity at specified times. We are particularly interested in the observations obtained by repeating the experiment several times.

In most cases, the outcomes will depend on chance and, therefore, cannot be predicted with certainty. If a chemist runs an analysis several times under the same conditions, he or she will obtain different measurements, indicating an element of chance in the experimental procedure. Even when a coin is tossed repeatedly, we cannot be certain that a given toss will result in a head. However, we know the entire set of possibilities for each toss.

Definition

The set of all possible outcomes of a statistical experiment is called the **sample space** and is represented by the symbol S.

Each outcome in a sample space is called a **basic outcome** or a **sample point**.

In the experiment of

- tossing a coin the basic outcomes are head (H) and tail (T), so $S = \{H, T\}$,
- rolling a die the basic outcomes are $1, 2, 3, 4, 5,$ and 6, so $S = \{1, 2, 3, 4, 5, 6\}$.

For any given experiment, we may be interested in the occurrence of certain events rather than in the occurrence of a specific element in the sample space. For instance, we may be interested in the event A that the outcome when a die is tossed is divisible by 3. This will occur if the outcome is an element of the subset $A = \{3, 6\}$.

Definition

An **event** is a subset of a sample space consisting of basic outcomes. The *null (impossible) event* represents the absence of a basic outcome and is denoted by $\emptyset$.

An event is said to occur if the random experiment results in one of the basic outcomes in that event.

In the experiment of rolling a die the basic outcomes are $1, 2, 3, 4, 5$, and 6, so $S = \{1, 2, 3, 4, 5, 6\}$. Let A be the event that the number on the die is even and B be the event that the number on the die is odd. Then $A = \{2, 4, 6\} \subset S$ and $B = \{1, 3, 5\} \subset S$. Event A is said to be occur if and only if the result of the experiment is one of 2, 4, or 6. Similarly, event B occurs if and only if the result of the experiment is 1, 3, or 5.

Intersection of Events, Union of Events, and Complement of an Event

Let A and B be two events in the sample space S.

- The intersection of A and B, denoted by $A \cap B$, is the set of all basic outcomes in S that belong to both A and B.

- The union of A and B, denoted by $A \cup B$, is the set of all basic outcomes in S that belong to at least one of A and B.

- The complement of A, denoted by $\bar{A}$ or A', is the set of all basic outcomes in S that doesn't belong to A.

Mutually Exclusive and Collectively Exhaustive Events

Definition

If the events A and B have no common basic outcomes, they are called **mutually exclusive** and $A \cap B = \emptyset$.

Note: For any event A, A and $\bar{A}$ are always mutually exclusive!

Mutually Exclusive and Collectively Exhaustive Events

Definition

If the events A and B have no common basic outcomes, they are called **mutually exclusive** and $A \cap B = \emptyset$.

Note: For any event A, A and $\bar{A}$ are always mutually exclusive!

Definition

Given the K events $E_1, E_2, \ldots, E_K$ in the sample space S, if $E_1 \cup E_2 \cup \cdots \cup E_K = S$, these K events are said to be *collectively exhaustive*.

Example. A die is rolled. If events A, B, and C are defined as:
A :"result is even",
B :"result is at least 4", and
C :"result is less than 6",
describe the events

a) $A \cap B$

b) $A \cup B \cup C$

c) A'

d) $A \cup B$

e) $A \cap B \cap C$

Counting Sample Points

One of the problems that the statistician must consider and attempt to evaluate is the element of chance associated with the occurrence of certain events when an experiment is performed. These problems belong in the field of probability.

In many cases, we shall be able to solve a probability problem by counting the number of points in the sample space without actually listing each element.

Basic Principle of Counting (The Multiplication Rule)

The fundamental principle of counting, often referred to as the **multiplication rule**:

If an operation can be performed in n_1 ways, and if for each of these ways a second operation can be performed in n_2 ways, then the two operations can be performed together in $n_1 n_2$ ways.

Example. A developer of a new subdivision offers prospective home buyers a choice of Tudor, rustic, colonial, and traditional exterior styling in ranch, two-story, and split-level floor plans. In how many different ways can a buyer order one of these homes?

Generalized Basic Principle of Counting (Generalized Multiplication Rule)

The multiplication rule may be extended to cover any number of operations:

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 $n_1 n_2 \cdots n_k$ ways.

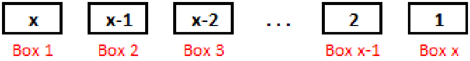
Example. In a medical study patients are classified according to their blood types (A, B, AB, and 0) and according to their blood pressure levels (low, normal, and high). In how many different ways a patient can be classified?

Example. How many different outcomes are possible if a coin is tossed twice and a die is rolled?

Permutations and Combinations

1. Number of orderings

We begin with the problem of ordering. Suppose that we have x objects that are to be **placed in order** in such a way that each object may be used **only once**.



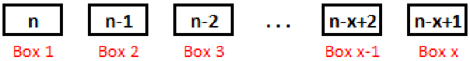
The total number of possible ways of arranging x objects **in order** is given by

$$x\,(x-1)\,(x-2)\cdots(2)\,(1) = x!,$$

where $x!$ is read "x factorial".

2. Permutations

Suppose that now we have n objects with which the x **ordered boxes** could be filled $(n > x)$ in such a way that each object may be used **only once**.



The total number of permutations of x object chosen from n, P_x^n, is the number of possible arrangements when x objects are to be **selected** from a total of n and **arranged in order**. This number is

$$P_x^n = n\,(n-1)\,(n-2)\cdots(n-x+2)\,(n-x+1) = \frac{n!}{(n-x)!}.$$

3. Combinations

Suppose that we are interested in the number of different ways that x can be **selected** from n (where no object may be chosen more than once) but the **order is not important**.

The number of combinations of x object chosen from n, C_x^n, is the number of possible selections that can be made. This number is

$$C_x^n = \frac{P_x^n}{x!} = \frac{\frac{n!}{(n-x)!}}{x!} = \frac{n!}{(n-x)!x!} = \binom{n}{x}.$$

PEARSON
IZMIR KATIP CELEBI UNIVERSITY

Example. In how many different ways can a person invite three of her eight closest friends to a party?

Example. In how many ways can we choose three letters from A,B,C,D

a) if the order is important?

b) if the order is not important?

Example. In how many different ways can the letters in UNUSUALLY be arranged?

Example. Consider a shuffled deck of 52 cards. In how many different ways can an Ace (**A**) be drawn?

Probability of an Event

Suppose that a random experiment is to be carried out and we want to determine the probability that a particular event will occur.

There are three definitions of probability:

1. Classical probability
2. Relative frequency probability
3. Subjective probability

In this course, we will focus only on the first one.

Classical Probability

Definition

Classical probability is the proportion of times that an event will occur (assuming that all outcomes in a sample space are equally likely to occur). The **probability of an event** A is

$$P(A) = \frac{n(A)}{n(S)},$$

where $n(A)$ is the number of outcomes that satisfy the condition of event A and $n(S)$ is the total number of outcomes in the sample space.

Note: Probability is measured over the range from 0 to 1!

Example. If two dice are rolled, what is the probability that the sum of the upturned faces will equal 7?

Solution.

$$\begin{aligned} S = \quad & \{(1,1),(1,2),(1,3),(1,4),(1,5),(1,6), \\ & (2,1),(2,2),(2,3),(2,4),(2,5),(2,6), \\ & (3,1),(3,2),(3,3),(3,4),(3,5),(3,6), \\ & (4,1),(4,2),(4,3),(4,4),(4,5),(4,6), \\ & (5,1),(5,2),(5,3),(5,4),(5,5),(5,6), \\ & (6,1),(6,2),(6,3),(6,4),(6,5),(6,6)\} \qquad \Rightarrow n(S) = 36 \end{aligned}$$

$$\begin{aligned} A \quad & = \text{sum of the upturned faces is } 7 \\ & = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\} \qquad \Rightarrow n(A) = 6 \end{aligned}$$

Therefore, $P(A) = \frac{n(A)}{n(S)} = \frac{6}{36} = \frac{1}{6} = 0.1667$.

Example. There are five laptops of which three are brand A and two are brand B. Two of them will be chosen at random.

a) Find the sample space.

b) Define the event E by "One brand A and one brand B laptops will be chosen." and list the elements of event E.

c) What is the probability that one brand A and one brand B laptops will be chosen? (Find $P(E)$.)

Example. Suppose that there are ten brand A, five brand B, and four brand C laptops and three of them will be chosen at random. What is the probability that two of them will be brand A and one will be brand C?

Example. A manager is available a pool of 8 employees who could be assigned to a project-monitoring task. 4 of the employees are women and 4 are men. 2 of the men are brothers. The manager is to make the assignment at random so that each of the 8 employees is equally likely to be chosen. Let A be the event that "Chosen employee is a man." and B be the event that "Chosen employee is one of the brothers."

a) Find $P(A)$.

b) Find $P(B)$.

c) Find $P(A \cap B)$.

Rules of Probability

The rules of probability are:

1. $P(S) = 1$
2. For any event $A \subset S$, $0 \leq P(A) \leq 1$
3. If $A_1, A_2, \ldots$ is a countable collection of mutually exclusive events, then

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

4. For any event $A \subset S$, $P(\bar{A}) = 1 - P(A)$
5. $P(\emptyset) = 0$
6. For any two events $A \subset S$ and $B \subset S$, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
7. For any three events $A \subset S$, $B \subset S$, and $C \subset S$, $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$

Example. The probability of A is 0.60, the probability of B is 0.45, and the probability of both is 0.30.

a) What is the probability of either A and B?

b) What are $P\left(\bar{A}\right)$ and $P\left(\bar{B}\right)$?

Example. A pair of dice is rolled. If A is the event that a total of 7 is rolled and B is the event that at least one die shows up 4, find the probabilities for A, B, $A \cap B$, and $A \cup B$.

Example. A corporation has just received new machinery that must be installed and checked before it becomes operational. The accompanying table shows a manager's probability assessment for the number of days required before the machinery becomes operational.

Number of days	3	4	5	6	7
Probability	0.08	0.24	0.41	0.20	0.07

Let A be the event "It will be more than four days before the machinery becomes operational." and let B be the event "It will be less than six days before the machinery becomes available."

a) $P(A) = ?$

b) $P(B) = ?$

c) $P(\bar{A}) = ?$

d) $P(A \cap B) = ?$

e) $P(A \cup B) = ?$

Example. Suppose the manufacturer's specifications for the length of a certain type of computer cable are 2000 ± 10 millimeters. In this industry, it is known that small cable is just as likely to be defective (not meeting specifications) as large cable. That is, the probability of randomly producing a cable with length exceeding 2010 millimeters is equal to the probability of producing a cable with length smaller than 1990 millimeters. The probability that the production procedure meets specifications is known to be 0.99.

a) What is the probability that a cable selected randomly is too large?

b) What is the probability that a randomly selected cable is larger than 1990 millimeters?

Conditional Probability, the Product Rule, and Independence

Conditional Probability

Suppose that two fair dice were rolled and we saw that one of them is 3. Under this condition, what is the probability of getting a total of 5?

Solution.
Since we know that one of the the dice is 3, we will be dealing with the restricted sample space

$$S_R = \{(1,3),(2,3),(3,1),(3,2),(3,3),(3,4),(3,5),(3,6),(4,3),(5,3),(6,3)\}$$

instead of the whole sample space $S = \{(1,1),(1,2),\ldots(6,6)\}$ consisting of 36 basic outcomes.

Therefore the desired event is $\{(2,3),(3,2)\}$ and the desired probability is $\frac{2}{11} = 0.1818$.

Definition

Let A and B be two events. The **conditional probability of event A, given that event B has occurred**, is denoted by $P(A|B)$ and is defined as

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \text{ provided that } P(B) > 0.$$

Similarly, the **conditional probability of event B, given that event A has occurred** is

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \text{ provided that } P(A) > 0.$$

Example. An international cargo company knows that 75% of its customers prefer shipments with SMS support while 80% prefer shipments with internet support. It is also known that 65% of the customers prefer both. What are the probabilities that

a) a customer who prefer SMS support will also prefer internet support?

b) a customer who prefer internet support will also prefer SMS support?

Example. The probability that there will be a shortage of cement is 0.28 and the probability that there will not be a shortage of cement and a construction job will be finished on time is 0.64. What is the probability that the construction job will be finished on time given that there will not be a shortage of cement?

Example. Consider an industrial process in the textile industry in which strips of a particular type of cloth are being produced. These strips can be defective in two ways, length and nature of texture. For the case of the latter, the process of identification is very complicated. It is known from historical information on the process that 10% of strips fail the length test, 5% fail the texture test, and only 0.8% fail both tests. If a strip is selected randomly from the process and a quick measurement identifies it as failing the length test, what is the probability that it is texture defective?

The Product (or Multiplicative) Rule

Definition

Let A and B be two events. Using the definitions of conditional probabilities $P(A|B)$ and $P(B|A)$, we have

$$P(A \cap B) = P(A|B)\, P(B)$$

$$P(A \cap B) = P(B|A)\, P(A).$$

Example. One bag contains 4 white balls and 3 black balls, and a second bag contains 3 white balls and 5 black balls. One ball is drawn from the first bag and placed unseen in the second bag. What is the probability that a ball now drawn from the second bag is black?

Statistical Independence

Definition

Let A and B be two events. A and B are said to be **statistically independent** if and only if

$$P(A \cap B) = P(A)\, P(B)\,.$$

More generally, the events $E_1, E_2, \ldots, E_K$ are **mutually statistically independent** if and only if

$$P(E_1 \cap E_2 \cap \cdots \cap E_K) = P(E_1)\, P(E_2) \cdots P(E_K)\,.$$

If A and B are (statistically) independent, then

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)\,P(B)}{P(B)} = P(A)\,, \text{ provided that } P(B) > 0$$

and

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A)\,P(B)}{P(A)} = P(B)\,, \text{ provided that } P(A) > 0.$$

Note: Do not confuse independent events with mutually exclusive events!

$$A \text{ and } B \text{ are independent} \iff P(A \cap B) = P(A)\,P(B)$$

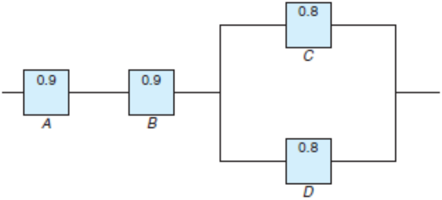
$$A \text{ and } B \text{ are mutually exclusive} \iff A \cap B = \emptyset \iff P(A \cap B) = 0.$$

Example. An electrical system consists of four components as illustrated in the figure. The system works if components A and B work and either of the components C or D works. The reliability (probability of working) of each component is also shown the figure. Find the probability that

a) the entire system works and

b) the component C does not work, given that the entire system works.

Assume that the four components work independently.



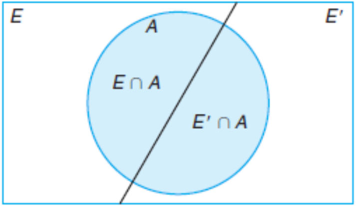
Bayes' Rule

Bayesian statistics is a collection of tools that is used in a special form of statistical inference which applies in the analysis of experimental data in many practical situations in science and engineering. **Bayes' rule** is one of the most important rules in probability theory. It is the foundation of Bayesian inference.

The Rule of Total Probability

We can write any event A as the union of the two mutually exclusive events $E \cap A$ and $E' \cap A$. Hence, $A = (E \cap A) \cup (E' \cap A)$ and

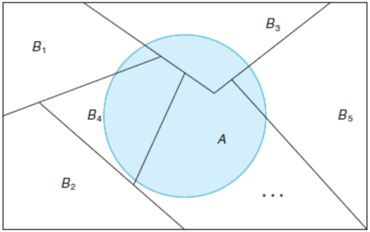
$$\begin{aligned} P(A) &= P\left((E \cap A) \cup (E' \cap A)\right) \\ &= P(E \cap A) + P(E' \cap A) \\ &= P(A|E)\,P(E) + P(A|E')\,P(E'). \end{aligned}$$



Theorem

If $B_1, B_2, \ldots, B_k$ are mutually exclusive and collectively exhaustive events such that $P(B_i) \neq 0$ $i = 1, 2, \ldots k$, then for any event A of S

$$P(A) = \sum_{i=1}^{k} P(B_i \cap A) = \sum_{i=1}^{k} P(A|B_i)\, P(B_i)$$



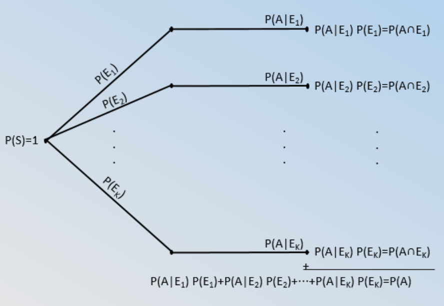
Bayes' Rule

Instead of asking for $P(A)$, suppose that we now consider the problem of finding the conditional probability $P(B_i|A)$.

Theorem

If $E_1, E_2, \ldots, E_K$ are mutually exclusive and collectively exhaustive events such that $P(E_i) \neq 0$ $i = 1, 2, \ldots K$, then for any event A in S such that $P(A) \neq 0$

$$P(E_i|A) = \frac{P(A|E_i)\,P(E_i)}{P(A|E_1)\,P(E_1) + P(A|E_2)\,P(E_2) + \cdots + P(A|E_K)\,P(E_K)}$$



Example. A hotel rents cars for its guests from three rental agencies. It is known that 25% are from agency X, 25% are from agency Y, and 50% are from agency Z. If 8% of the cars from agency X, 6% from agency Y, and 15% from agency Z need tune-ups, what is the probability that a car needing a tune-up come from agency Y?

Example. A life insurance salesman finds that, of all the sales he makes, 70% are to people who already own policies. He also finds that, of all contacts for which no sale is made, 50% already own life insurance policies. Furthermore, 40% of all contacts result in sales. What is the probability that a sale will be made to a contact who already owns a policy?

Example. In a large city, 8% of the inhabitants have contracted a particular disease. A test for this disease is positive in 80% of people who have the disease and is negative in 80% of people who do not have the disease. What is the probability that a person for whom the test result is positive has the disease?

Example. A record-store owner assesses customers entering the store as high school age, college age, or older, and finds that of all customers 30%, 50%, and 20%, respectively, fall into these categories. The owner also found that purchases were made by 20% of high school age customers, by 60% of college age customers, and by 80% of older customers.

a) What is the probability that a randomly chosen customer entering the store will make a purchase?

b) If a randomly chosen customer makes a purchase, what is the probability that this customer is high school age?

Example. A restaurant manager classifies customers as regular, occasional, or new, and finds that of all customers 50%, 40%, and 10%, respectively, fall into these categories. The manager found that wine was ordered by 70% of the regular customers, by 50% of the occasional customers, and by 30% of the new customers.

a) What is the probability that a randomly chosen customer orders wine?

b) If wine is ordered, what is the probability that the person ordering is a regular customer?

c) If wine is ordered, what is the probability that the person ordering is an occasional customer?