STA255: Statistical Theory

Chapter 2: Probability

Summer 2017

A Review of Set Notation

- For probability theory, we need some basic concepts of set theory.
- We will use capital letters, A, B, C, . . . to denote sets.
- If the elements in the set A are a_1, a_2 and a_3, we will write:

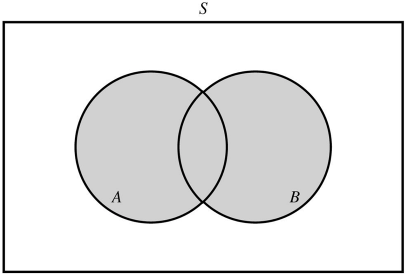
$$A = \{a_1, a_2, a_3\}.$$

- Let S denotes the set of all elements under consideration. Then S is called the universal set.
- We say that A is a subset of B, or A is contained in B (denoted $A \subset B$), if every point in A is also in B.
- The null set (empty set), denoted by ϕ, is the set consisting of no points. Thus, ϕ is a subset of every set.

Sets Operations

- The union of A and B, denoted by $A \cup B$, is the set of all points in A or B or both.

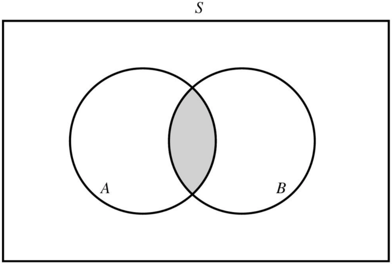
$$A \cup B = \{x \in S : x \in A \text{ or } x \in B\}.$$



Sets Operations

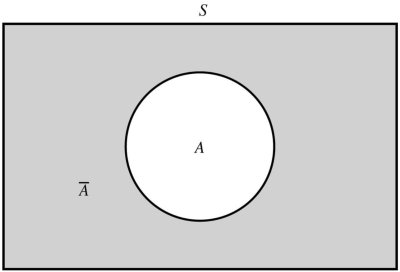
- The intersection of A and B, denoted by $A \cap B$, is the set of all points in both A and B.

$$A \cap B = \{x \in S : x \in A \;\text{ and }\; x \in B\}.$$



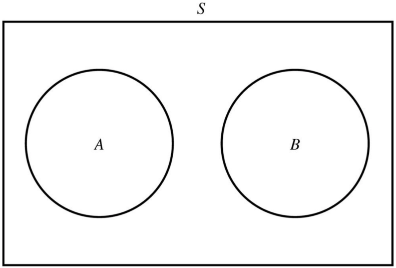
Sets Operations

- The complement of a set A, denoted by $\overline{A}$, is the set of all points in S that are not contained in A.
- Note: $A \cup \overline{A} = S$.



Sets Operations

- When A and B have no elements in common (i.e. $A \cap B = \phi$), they are said to be mutually exclusive or disjoint events.



Sets Operations

- Distributive Laws:
  - $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
  - $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

Sets Operations

- De Morgan's Law:
  - $\overline{A \cap B} = \overline{A} \cup \overline{B}$
  - $\overline{A \cup B} = \overline{A} \cap \overline{B}$

Proof:

More Properties

(1) If $A \subset B$, then $A \cap B = A$ and $A \cup B = B$.

(2) $A \cap \phi = \phi$ and $A \cup \phi = A$.

More Properties

(3) $\overline{\overline{A}} = A$.

(4) A but not $B = A - B = A \cap \overline{B}$.

Basic Concepts: Random Experiments

- An experiment is a process that generates data.
- A random experiment is an experiment in which:
    1. All possible outcomes of the experiment are known in advance.
    2. Any performance of the experiment results in an outcome that is not known in advance.
    3. The experiment can be repeated under identical conditions.
- Examples:
- Tossing a coin once or several times.
- Rolling a die once.
- Examine fuse for a defeat.

Basic Concepts: Sample Space

- The Sample Space of an experiment, denoted by S, is the set of all possible outcomes.
- Sample spaces are either:
  - discrete (contains a finite number of elements, or an infinite but countable number of elements) or
  - continuous (an infinite number of sample points constituting a continuum).

More details will be given later when talk about **Random Variables**.

Basic Concepts: Sample Space

- Examples:
  1. Rolling a die once
  $$S = \{1, 2, 3, 4, 5, 6\}.$$
  Here the outcomes correspond to the side that turns up.
  2. Examine a fuse for a defect (N: not defective, D: defective)
  $$S = \{N, D\}$$
  3. Examine two fuses in sequence and note the outcome
  $$S = \{NN, ND, DN, DD\}$$
  4. Examine each fuse as it comes off the assembly line until the first defective fuse is found. Note the number examined
  $$S = \{1, 2, 3, \ldots\}$$

Basic Concepts: Events

- An event is any subset of the sample space.
- An event is said to be simple if it consists of exactly one outcome and compound if it consists of more than one outcome.
- Example: Consider the experiment of rolling two dice and define the following events:
  1. $A =$ The sum of the two numbers is 5.
  $$A = \{(1,4),(4,1),(2,3),(3,2)\}.$$
  2. $B =$ Same number on both dice.
  $$C = \{(1,1),(2,2),(3,3),(4,4),(5,5),(6,6)\}.$$

Probability Axioms

- Given an experiment and a sample space S, the objective of probability is to assign to each event A a number $P(A)$, called the probability of the event A, which will give a precise measure of the chance that A will occur.
- Probability Axioms:
  1. Axiom 1: For any event A, $P(A) \geq 0$.
  2. Axiom 2: $P(S) = 1$.
  3. Axiom 3: If $A_1, A_2, \ldots$ is an infinite collection of mutually exclusive events (i.e. $A_i \cap A_j = \phi$, for $i \neq j$), then

$$P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots = \sum_{i=1}^{\infty} P(A_i).$$

Proposition

Proposition

$P(\phi) = 0$

Assume $P(\phi) > 0$. For any event A, we have $A \cap \phi = \phi$ and

$$A = A \cup \phi \cup \phi \cup \phi \dots$$

$$P(A) = P(A \cup \phi \cup \phi \cup \phi \cup \dots) \quad \text{from Axiom 3}$$

$$P(A) = P(A) + P(\phi) + P(\phi) + P(\phi) + \dots > P(A)$$

Contradiction! Thus $P(\phi) = 0$

Proposition

Proposition

If $A_1, A_2, \cdots, A_n$ is a finite collection of mutually exclusive events, then

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n) = \sum_{i=1}^{n} P(A_i).$$

$$\begin{aligned}
P(A_1 \cup A_2 \cup \cdots \cup A_n) &= P(A_1 \cup A_2 \cup \cdots \cup A_n \cup \phi \cup \phi \cup \ldots) \\
&= P(A_1) + P(A_2) + \cdots + P(A_n) + P(\phi) + P(\phi) + \ldots \\
&= P(A_1) + P(A_2) + \cdots + P(A_n) \\
&= \sum_{i=1}^{n} P(A_i)
\end{aligned}$$

Example: #2.22

If A and B are events and $B \subset A$. Show that

$$P(A) = P(B) + P(A \cap \overline{B}).$$

Solution:

Computing Probabilities of Events

- Equally likely outcomes (classical probability): We say the outcomes of sample space with N elements/objects:

$$S = \{E_1, E_2, \cdots, E_N\}$$

  are equally likely, if the probability assigned to each element is the same value, i.e. $P(E_i) = \frac{1}{N}$.

- If the sample space outcomes are equally likely to occur, then:

$$\begin{aligned} P(A) &= \frac{\text{Number of outcomes in } A}{\text{Number of outcomes in } S} \\ &= \frac{n(A)}{n(S)}. \end{aligned}$$

Sample Point Method

1. Define the experiment and describe a sample space, S.
2. List all the simple events
3. Assign probabilities to the sample points in S;

$$P(E_i) \geq 0 \text{ and } \sum_i P(E_i) = 1$$

4. Define the event A as a collection of sample points
5. Calculate $P(A)$ by summing the probabilities of sample points in A.

See the example 2.3 on page 37

Example: #2.18

Suppose two balanced coins are tossed and the upper faces are observed.

(a) List the sample points for this experiment.

(b) Assign a reasonable probability to each sample point. (Are the sample points equally likely?)

(c) Let A denote the event that exactly one head is observed and B the event that at least one head is observed. List the sample points in both A and B.

(d) From your answer to part (c), find $P(A)$, $P(B)$, $P(A \cap B)$, $P(A \cup B)$, and $P(\overline{A} \cup B)$.

Solution:

Example: #2.18

Tools for Counting Sample Points

- Recall: If the sample space outcomes are equally likely to occur, then:

$$P(A) = \frac{\text{Number of outcomes in } A}{\text{Number of outcomes in } S} = \frac{n(A)}{n(S)}.$$

- Example:
  - In the die-rolling experiment, $S = \{1, 2, 3, 4, 5, 6\}$.
  - Let $A =$ even numbers. That is, $A = \{2, 4, 6\}$.
  - $P(A) = \frac{n(A)}{n(S)} = \frac{3}{6} = 0.5$.

Tools for Counting Sample Points

- It is important to be able to count the number of possible outcomes in an experiment.
- Counting the outcomes of an experiment can easily become quiet large.
- Counting the outcomes is difficult, unless we know counting rules.
- We will study the following counting principles:
  - Addition and Product Rules
  - Permutations (without replacement and order is important)
  - Combinations (without replacement and order is NOT important)

Counting Principles: Addition and Product Rules

- The Addition Principle: If a choice from Group I can be made in n ways and a choice from Group II can be made in m ways, then the number of choices possible from Group I OR Group II is $n + m$.
- Example: Enrollment in the course Principles of probability consists of: 28 statistics majors, of whom 10 are males, and 53 math majors, of whom 4 are males. One of the enrolled students is selected at random. The number of ways to select a male student is $10 + 4 = 14$.

Counting Principles: Addition and Product Rules

- The Product Principle: If a task involves two steps and the first step can be completed in n ways AND the second step in m ways, then there are $n \times m$ ways to complete the task.
- Example: The number of ways to select one math and one statistics student is

$$53 \times 28 = 1484.$$

- The general product rule: If an experiment can be completed in k stages and stage i has n_i outcomes then the experiment has

$$n_1 \times n_2 \times \ldots n_k \text{ outcomes}$$

Examples

- The door on the computer center has a lock which has five buttons numbered from 1 to 5. The combination of numbers that opens the lock is a sequence of five numbers and is reset every week.

  (a) How many combinations are possible if a button can only be used once?
  Number of ways $= 5 \times 4 \times 3 \times 2 \times 1 = 120.$

  (b) How many combinations are possible if there is no restriction on the number of times a button can be used?
  Number of ways $= 5 \times 5 \times 5 \times 5 \times 5 = 5^5 = 3125.$

Partition

- The number of ways of partitioning n distinct objects into k distinct groups containing $n_1, n_2, \cdots, n_k$ objects, respectively, where each object appears in exactly one group and $n_1 + n_2 + \ldots + n_k = n$, is

$$\binom{n}{n_1 \; n_2 \; \ldots \; n_k} = \frac{n!}{n_1! \, n_2! \, \ldots \, n_k!}$$

- Example: If I have 3 A's,4 B's and 1 C, how many ways to arrange all 8 letters? (without replacement)

  Number of ways $= \frac{8!}{3!4!1!} = 280$.

Permutations of size r from n letters

- An ordered arrangement of r distinct objects is called a permutation.
- $P_{n,r} = P_r^n =$ number of ways of ordering n distinct objects taken r at a time.
- Example: Write all the permutations of size 2 from 4 letters a, b, c, d.
  Answer: $ab, ac, ad, ba, bc, bd, ca, cb, cd, da, db, dc$.
- Theorem: $P_r^n = \frac{n!}{(n-r)!} = n \times (n-1) \cdots \times (n-r+1)$.
- Note: To write a permutation we should not repeat any object and the order that an object appears is important. For example *ab* and *ba* are different.

Examples

- Among a group of 10 drivers 3 are to be selected to go to 3 different locations. In how many ways can this be done?

$$P_{10,3} = \frac{10!}{(10-3)!} = 10 \times 9 \times 8 = 720$$

Counting Principles: Combinations

- The number of combinations of n objects taken r at a time is the number of subsets, each of size r, that can be formed from the n objects. This number will be denoted by

$$\binom{n}{r} = C_r^n.$$

- A combination is a collection of elements whose order does not matter.
- Theorem: We have

$$\binom{n}{r} = C_{n,r} = \frac{n!}{r!(n-r)!}.$$

Counting Principles: Combinations

- Example: Write all the combinations of 2 letters from the letters a, b, c, d.

  Answer: ab, ac, ad, bc, bd, cd.

- Notice that we did not include ba when ab is included. That is, ab and ba are assumed to be identical combinations.

Counting Principles: Combinations

- Facts:
  1. $r!C_r^n = \mathrm{P}_r^n$.
  2. Binomial Theorem: $(a+b)^n = \sum_{r=0}^{n} \binom{n}{r} a^r b^{n-r}$
  3. If $a = b = 1$ in (2), then $\sum_{r=0}^{n} \binom{n}{r} = 2^n$.
  4. $\binom{n}{r} = \binom{n}{n-r}$

Examples and Applications in Probability

1. There are 20 computers in a store. Among them 15 are brand new and 5 are refurbished. 6 computers are purchased for a student lab. From the first look, they look indistinguishable, so the 6 computers are selected at random. Compute the probability that among the chosen computers, 2 are refurbished.

   Solution:

$$P(2 \text{ are refurbished}) = \frac{C_2^5\, C_4^{15}}{C_6^{20}} = 0.3522.$$

Conditional Probability

- The conditional probability of an event A, given that event B has occurred is given by
$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$
provided $P(B) > 0$.

- Note: By cross-multiplying both sides, we get "the multiplication rule":
$$P(A \cap B) = P(A|B)P(B)$$

- It is also true that
$$P(A \cap B) = P(B|A)P(A)$$

This is most useful when the experiment consists of two stages and events A and B pertain two outcomes of stages 1,2, respectively.

Example

The probability that a regularly scheduled flight departs on time is $P(D) = 0.83$; the probability that it arrives on time is $P(A) = 0.82$; and the probability that it departs and arrives on time is 0.78.

(a) Find the probability that a plane arrives on time given that it departed on time.

$$P(A|D) = \frac{P(A \cap D)}{P(D)} = \frac{0.78}{0.83} = 0.94$$

(b) Find the probability that a plane departed on time given that it has arrived on time.

$$P(D|A) = \frac{P(A \cap D)}{P(A)} = \frac{0.78}{0.82} = 0.95.$$

(c) Find the probability that a plane that it arrives on time, given that it did not depart on time.

$$P(A|\overline{D}) = \frac{P(A \cap \overline{D})}{P(\overline{D})} = \frac{0.82 - 0.78}{0.17} = 0.24$$

Independence

- Two events A and B are said to be independent if any one of the following holds:
  1. $P(A|B) = P(A)$ or, equivalently,
  2. $P(B|A) = P(B)$ or, equivalently,
  3. $P(A \cap B) = P(A)P(B)$.
- Events that are not independent are often said to be dependent.
- In general, the idea behind independence is that two events are independent if knowledge about one event occurring gives us no information about whether the other event occurred.

More about Mutually Exclusive and Independence

- **Independence** does not imply that the sets do not intersect.
- **Mutually Exclusive** is a property of sets: $A \cap B = \phi$.
- Independence is a property of probability: $P(A \cap B) = P(A)P(B)$.
- The following results show how divergent are the two concepts are: If A and B are two events such that $P(A) > 0$ and $P(B) > 0$, then
  1. If A and B are independent, then they CANNOT be mutually exclusive. (Give a reason)
  2. If A and B are mutually exclusive, then they CANNOT be independent. (Give a reason)

Addition Rule

Theorem

For any two events A and B,

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

Complement Rule

Theorem

For any event A,

$$P(\overline{A}) = 1 - P(A).$$

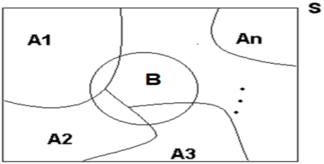
Independence

- Example: If A and B are independent, show that A and $\overline{B}$, $\overline{A}$ and B, and $\overline{A}$ and $\overline{B}$ are independent as well.

Law of Total Probability

- Partition of Sample Space: Let $A_1, A_2, \cdots, A_n$ be subsets of a sample space S. If $A_i \cap A_j = \phi$ for all $i \neq j$,and $A_1 \cup A_2 \cup \cdots A_n = S$, then the sequence $A_1, A_2, \cdots, A_n$ is called a partition.
- Law of Total Probability: Let $A_1, A_2, \cdots, A_n$ constitute a partition of the sample space S such that $P(A_i) > 0$ for $i = 1, 2, \ldots, n$, then for any event B in S such that $P(B) \neq 0$,

$$\begin{aligned} P(B) &= P(B \cap A_1) + \cdots + P(B \cap A_n) \\ &= P(B|A_1)P(A_1) + \cdots + P(B|A_n)P(A_n) \\ &= \sum_{i=1}^{n} P(B|A_i)P(A_i) \end{aligned}$$



Bayes' Theorem

- Bayes' Theorem (Bayes' rule): Let $A_1, A_2, \cdots, A_n$ constitute a partition of the sample space S such that $P(A_i) > 0$ for $i = 1, 2, \ldots, n$, then for any event B in $S such that$P(B) $\neq 0$,

$$P(A_k|B) = \frac{P(B \cap A_k)}{P(B)} = \frac{P(B|A_k)P(A_k)}{P(B)} = \frac{P(B|A_k)P(A_k)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)}$$

- Bayes'Theorem, special case:

$$P(A|B) = \frac{P(B \cap A)}{P(B)} = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\overline{A})P(\overline{A})}$$

Example

In a certain assembly plant, three machines, B_1, B_2, and B_3, make 30%, 45%, and 25%, respectively, of the products. It is known from past experience that 2%, 3%, and 2% of the products made by each machine, respectively, are defective.

(a) Suppose that a finished product is randomly selected. What is the prob- ability that it is defective?

(b) If a product was chosen randomly and found to be defective, what is the probability that it was made by machine B_3?

Solution: