Chapter 2

- Probability is the measure of one's belief that a future (random) event will occur.
- A random event is one whose occurrence cannot be predicted with certainty.
- However, we can often understand the long-run relative frequency (proportion) of times the event will occur in a long series of trials.

Examples of random events:

- Toss of a coin
- Gain/loss of the Dow Jones stock index
- Size of randomly selected item from a certain population.

Axiomatic method

Defn: A random experiment is the process by which an observation is made.
Example: We roll a die and record the outcome.

Defn: An outcome (also called a sample point) is the particular result of an experiment.

A probability is assigned to each outcome. The probability of more complex events is computed using axiomatic rules.

Probability and Inference

If die is balanced, $\mathbb{P}(1) =$
Suppose out of 50 tosses, we obtain no "1"s.
Is it impossible?
Is this unlikely?
Our conclusion is an inference based on a probability.

Set Notation

We use capital letters to denote sets of points (points are outcomes of an experiment in our case).

Special Sets: $S =$ set of all points (outcomes). It is called probability space or sample space.
$\emptyset =$ empty set (set with no points).
Example: $S = \{1, 2, 3, 4, 5, 6\}$ for a toss of a die.

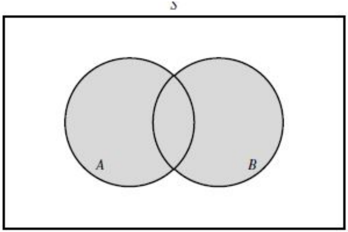
Defn: A is a subset of B (denoted $A \subset B$) if every point in A is also in B.
Example: $\{1, 4\} \subset \{1, 2, 3, 4\}$.

Union and intersection of sets

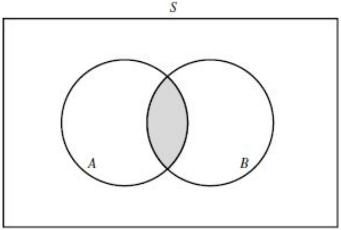
Defn: The union of A and B (denoted $A \cup B$) is the set of all points in A or in B or in both.
Defn: The intersection of A and B (denoted $A \cap B$ or AB) is the set of all points that are in A and in B simultaneously.

A Venn Diagram can graphically display simple sets.

$A \cup B$



$A \cap B$

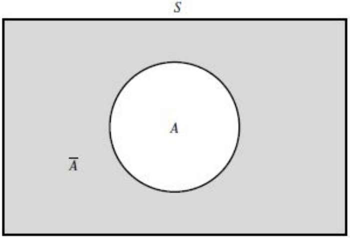


Complement set and mutually exclusive sets

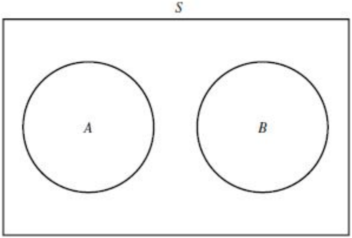
Defn: A complement of A (denoted $\overline{A}$) is the set of points that are in S but not in A.

Defn: Sets A and B are mutually exclusive (or disjoint) if $A \cap B = \emptyset$. (Mutually exclusive sets have no points in common.)

$\overline{A}$



Disjoint A and B



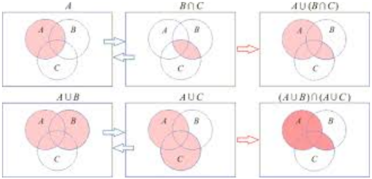
Die-rolling example

$S = \{1, 2, 3, 4, 5, 6\}$
Let $A = \{\text{roll number is less than 3}\}$.
Let $B = \{\text{roll odd number}\}$.
Let $C = \{\text{roll even number}\}$.

$A \cap B = AB =$
$A \cap C = AC =$
$A \cup C =$
$\overline{A} =$

Which are mutually exclusive?
A and B? A and C? B and C?

Important laws



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

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

Section 2.4. Discrete Probability Models

Recall:

- An experiment is the process by which an observation is made.
- A sample point (or outcome) is the particular result of an experiment.
- The sample space (or probability space) is the set of all possible sample points.

Defn: A discrete sample space contains a finite or countable number of sample points.

Examples:

1. Roll a die:
   $S =$
2. Flip a coin until the first "head" and record the number of "tails" that occurred.
   $S =$
3. Randomly select an American household and record the number of TV sets.
   $S =$

Events

Defn: An event is a collection of sample points (an event is a subset of S).
A simple event correspond to exactly one sample point.

Example. Roll a die.
Event $A = \{\text{roll an odd number}\}$
Event $B = \{\text{roll a 2}\}$

Example. Flip a coin three times.
Event $C =$"get heads on all three flips"
$C =$

Event $D =$"get head before the third flip"
$D =$

Probability Axioms (Kolmogorov)

Let S be a sample space. We assign $\mathbb{P}(A)$, the probability of A, to each event $A \subset S$, in such a way that the following axioms are satisfied:

Axiom 1: $\mathbb{P}(A) \geq 0$.

Axiom 2: $\mathbb{P}(S) = 1$.

Axiom 3: If $A_1, A_2, A_3, \ldots$ is a sequence of pairwise mutually exclusive events, then

$$\mathbb{P}(A_1 \cup A_2 \cup A_3 \cup \ldots) = \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

Corollary 1: $\mathbb{P}(\emptyset) = 0$.
Proof: If $\mathbb{P}(\emptyset) > 0$, then by Axiom 3,
$\mathbb{P}(S) = \mathbb{P}(S \cup \emptyset \cup \emptyset \cup \ldots) = \mathbb{P}(S) + \sum_{i=2}^{\infty} \mathbb{P}(\emptyset) = \infty$,
which contradicts Axiom 2. $\square$

Probability of a finite union of disjoint events

Corollary 2: If $A_1, A_2, A_3, \ldots, A_n$ are pairwise mutually exclusive events, then

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

Proof: $\mathbb{P}(A_1 \cup A_2 \cup \ldots \cup A_n) = \mathbb{P}(A_1 \cup A_2 \cup \ldots \cup A_n \cup \emptyset \cup \ldots)$. The result follows by application of Axiom 3 and Corollary 1. $\square$

Note: If we know the probability of all simple events, then we can compute the probability of all events by Axiom 3 or Corollary 2.

Example: $\mathbb{P}(1) = 1/4$, $\mathbb{P}(3) = 1/4$, $\mathbb{P}(5) = 1/4$.
$\mathbb{P}(\{1, 3, 5\}) =$

Probability Rules following from the Axioms

Complement Rule: $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$.
Proof: $\mathbb{P}(A \cup \overline{A}) = \mathbb{P}(S) = 1$. On the other hand, by Corollary 2, $\mathbb{P}(A \cup \overline{A}) = \mathbb{P}(A) + \mathbb{P}(\overline{A})$. Hence, $\mathbb{P}(A) + \mathbb{P}(\overline{A}) = 1$. $\square$

Monotonicity: If $A \subset B$, then $\mathbb{P}(A) \leq \mathbb{P}(B)$.
Proof: B is the union of two disjoint events, A and $B \cap \overline{A}$. Hence, $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \cap \overline{A}) \geq \mathbb{P}(A)$ by Axiom 1. $\square$

Corollary: For any $A \subset S$, $\mathbb{P}(A) \leq \mathbb{P}(S) = 1$.

Additive rule of probability

Theorem $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

Proof:

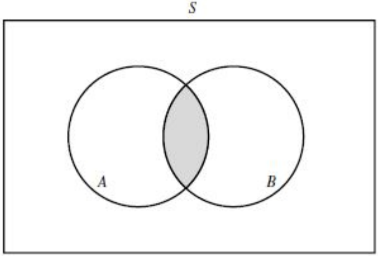
$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap \overline{B})$

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

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

Hence,

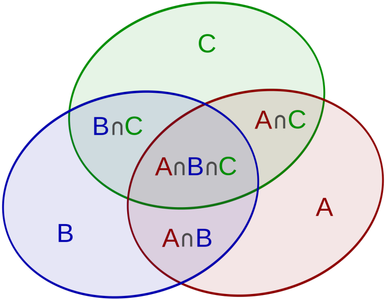
$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$



$\square$

Inclusion-Exclusion formulas

$\mathbb{P}(A \cup B \cup C) =$
$\mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C).$



In general, an extended additive law holds for any n-events A_1, A_2, $\ldots$,A_n:
$\mathbb{P}(A_1 \cup A_2 \cup \ldots \cup A_n) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \ldots + \mathbb{P}(A_n) - \mathbb{P}(A_1 \cap A_2) - \ldots - \mathbb{P}(A_{n-1} \cap A_n) + \mathbb{P}(A_1 \cap A_2 \cap A_3) + \ldots \pm \mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_n)$

Sample-point method

- List the points corresponding to an event A.
- Determine an appropriate probability for each sample point.
- Sum the probabilities of the sample points in A.

This will give the probability of the event A.

Example: Toss three coins, each weighted such that "heads" is twice as likely as "tails". What is the probability of the event
$A = \{\text{we got at least 2 "heads"}\}$
Sample Space:
HHH $\rightarrow$ 8/27
HHT $\rightarrow$ 4/27
HTH $\rightarrow$ 4/27
HTT $\rightarrow$ 2/27
THH $\rightarrow$ 4/27
THT $\rightarrow$ 2/27
TTH $\rightarrow$ 2/27
TTT $\rightarrow$ 1/27

The case of equally likely outcomes

When all sample points in S are equally likely, finding $\mathbb{P}(A)$ is easier. In this case:

$$P(A) = \frac{\text{Number of sample points in } A}{\text{Total number of sample points}}$$

Example: same experiment, except all three coins are “fair”.

Sample Space:
HHH
HHT
HTH
HTT
THH
THT
TTH
TTT

Example

Roll two fair dice. Consider the event “The sum of the two dice is 7”. What is the probability of this event?

The sample space is $\{(1,1),(1,2),\ldots,(6,5),(6,6)\}$
- 36 ordered pairs.

Example: Choice of pivot in Quicksort

The problem is to sort an array, for example, $3, 8, 2, 5, 1, 4, 7, 6$.
$(n = 8)$

The basic idea is to choose a pivot element and partition around the pivot. For a pivot "=3", we obtain $2, 1, 3, 6, 7, 4, 5, 8$. This operation takes about n swaps where n is the length of the array.

Then recourse on the two new subarrays. For "good" choice of pivots this will give around $n \log_2 n$ operations.

Consider the event "A randomly chosen pivot gives the 25-75 split or better". What is the probability of this event?

The sample space is $\{1, ..., n\}$.

$S = \{(n/4 + 1)\text{-th smallest element}, \ldots, 3n/4\text{-th smallest element}\}$

Probability:

Tools for Counting Sample Points

Results from combinatorial analysis are useful for counting sample points.

Result: Consider selecting an element from a set having m elements and selecting another from a set having n elements. Then there are mn possible pairs that could be selected.

Note: This multiplication rule extends to any number of sets.

Exam[le We roll a 6-sided die, a 4-sided die, and s 20-sided die.

The total number of possibilities is $6 \times 4 \times 20$.

Permutations

Defn: A permutation is an ordered arrangement of a specific number of distinct objects. (Order matters !)
We denote the number of possible permutations of n distinct objects, taken r at a time, as P_r^n.

A formula for P_r^n: In how many ways can we fill r positions with n distinct objects ?

$$P_r^n = n \times (n-1) \times (n-2) \times \ldots \times (n-r+1) = \frac{n!}{(n-r)!},$$

where $n!$ denote the factorial of n:
$n! = n \times (n-1) \times (n-2) \ldots \times 2 \times 1.$

Example: There are 24 members of a student club. Four members will be chosen to be president, VP, secretary, and treasurer. How many different arrangements of officers are possible?

Permutations when not all objects are distinct

Example: How many permutations of the letters of the name "BOBBY" are there?

Theorem: The number of permutations of n objects, in which n_1 are in the first group, n_2 is in the second group, $\ldots$, n_k is in the k-th group, and $n_1 + n_2 + \ldots + n_k = n$, is:

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

In the previous example:

$$\binom{5}{3, 1, 1} = \frac{5!}{3!1!1!} = \frac{5!}{3!}$$

Note: If there are only two groups, a shortcut notation is often used:

$$\binom{n}{k, n-k} = \binom{n}{k}$$

Example

Example Consider a sequence of 10 coins, of which 2 are pennies, 5 are nickels and 3 are dimes. How many distinguishable arrangements are possible?

Partitioning

Example 10 children are to be divided into 3 groups: 3 children will be in classes A and B, and 4 children in class C. How many ways to do it?

Theorem: The number of ways to partition n objects in k distinct groups, containing $n_1, n_2, \ldots, n_k$ objects, respectively, where each object is in exactly one group and $n_1 + \ldots + n_k = n$, is

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

Note: The numbers $\binom{n}{n_1, n_2, \ldots, n_k}$ are also called multinomial coefficients.

Number of combinations

Example: How many ways to select a 3-person commission from 13 faculty members?

Theorem The number of unordered subsets of size r chosen from n objects is

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

Note: This number is also called the number of combinations and sometimes denoted C_r^n. The numbers $C_r^n = \binom{n}{r}$ are also called binomial coefficients.

Examples

In poker, a full house is a hand that contains 3 cards of one rank and 2 cards of another rank. What is the probability of being dealt a full house?

A flush is a hand that contains 5 cards of the same suit. What is the probability of being dealt a flush?

Which probability is larger?

Quiz

A fair coin is tossed 4 times. What is the probability that we get exactly two heads?

Quiz

If we toss 10 fair coins, what is the probability that we obtain two or more "heads"?

(A) About 0.9

(B) About 0.99

(C) About 0.999

(D) About 0.9999

(E) About 0.99999

Quiz

A group of people consists of 3 girls and 4 boys.
In how many ways they can select a committee that consists of 2 girls and 2 boys?

Event-Composition Method

Identify the event of interest and use the laws of probability to find its probability.

Example of a probability law: $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cup B)$.

Many probability laws involve the concepts of conditional probability and independence.

Conditional probability

Consider randomly selecting a 40-year-old person.

Let event A=“the person will contract lung cancer in the next two decades,“ and event B = “the person is a smoker“.
What is the probability of the event A given that we know that event B is true?

This is a conditional probability, which we denote $\mathbb{P}(A|B)$. It may be different from the unconditional probability $\mathbb{P}(A)$.

Defn: Given that $\mathbb{P}(B) > 0$, the conditional probability of A given B is defined by

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

Example We roll a fair die. Define A=“roll a 6“ and B=“roll an even number”. What is $\mathbb{P}(A|B)$?
$P(A \cap B) = P(A) =$
$P(B) =$
$P(A|B) =$

Quiz

Suppose you roll two fair dice. What is the probability that at least one die is a 1, given that the sum of the two dice is 7?

(A) 1/36
(B) 1/6
(C) 1/3
(D) 1/2

Properties of conditional probability

The conditional probability satisfy the three Kolmogorov's axioms with the sample space $S = B$. Hence it satisfies all probability laws previously derived.

For example, $\mathbb{P}(\overline{A}|B) = 1 - \mathbb{P}(A|B)$.

Independence

Defn: Two events A and B are independent if (and only if) any one of these conditions is true:

1. $\mathbb{P}(A|B) = \mathbb{P}(A)$
2. $\mathbb{P}(B|A) = \mathbb{P}(B)$
3. $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$

If any of these conditions is false, then A and B are called dependent.

Example: We roll a fair die. Define A="roll a 6", B="roll an even number", and C="roll a number greater than 4".

Are A and C independent?

Are B and C independent?

Quiz

Suppose you roll two fair dice.
Consider the events: “the first die is a 1” and “the sum of the two dice is 7”. Are they independent?

(A) Yes
(B) No

Multiplicative law of probability

Theorem The probability of intersection of two events, A and B, is

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B) = \mathbb{P}(B|A)\mathbb{P}(A).$$

For independent events A and B,

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

Example: Suppose a basketball player makes 70% of his initial free throws. On his second attempt, he has an 80% success rate if he made the first and a 60% success rate if he missed the first.

What is the probability he makes both of a pair of free throws?

What is the probability he misses both free throws?

Birthday Problem

What is the probability that in a set of n randomly chosen people some pair of them will have the same birthday.

What is an approximate formula? (illustration in figure)

What is the number of people that is needed to make the probability greater than $p = 90\%$

http://www.cornell.edu/video/the-tonight-show-with-johnny-carson-feb-8-1980-excerpt

The law of Total Probability

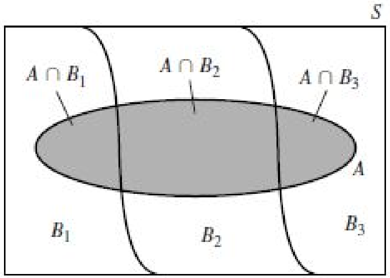
Defn: For some positive integer k, a collection of sets $\{B_1, B_2, \ldots, B_k\}$ forms a partition of S if:

1. $B_i \cap B_j = \emptyset$ if $i \neq j$.
2. $S = B_1 \cup B_2 \cup \ldots \cup B_k$.

Theorem: (Law of Total Probability)
If $\{B_1, B_2, \ldots, B_k\}$ is a partition of S such that $\mathbb{P}(B_i) > 0$, $i = 1, \ldots, k$, then for any event A,

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

Proof of the law of total probability



Proof: $A = (A \cap B_1) \cup (A \cap B_2) \cup \ldots (A \cap B_n)$ and the terms in this union are disjoint.
By additive law, $\mathbb{P}(A) = \mathbb{P}(A \cap B_1) + \mathbb{P}(A \cap B_2) + \ldots + \mathbb{P}(A \cap B_n)$.

By applying the multiplicative law to each term in this sum, we get:
$\mathbb{P}(A) = \mathbb{P}(A|B_1)\mathbb{P}(B_1) + \mathbb{P}(A|B_2)\mathbb{P}(B_2) + \ldots + \mathbb{P}(A|B_n)\mathbb{P}(B_n)$. $\square$

Example

An urn contains 5 green and 8 white balls. A ball is drawn and its color is noted. Then it is returned to the urn together with two new balls of the same color. Then the procedure is repeated.

Let G_2 denotes the event that the ball drawn in the second round is green. What is its probability?

What is $\mathbb{P}(G_3)$?

Difference between $\mathbb{P}(A|B)$ and $\mathbb{P}(B|A)$

Example:
Toss a coin 5 times. Let A = “first toss is heads” and let B = “all 5 tosses are heads”.

Then $\mathbb{P}(A|B) = 1$ but $\mathbb{P}(B|A) = 1/16$.

Bayes' Rule

Theorem: If $\{B_1, B_2, \ldots, B_k\}$ is a partition of S such that $\mathbb{P}(B_i) > 0$, $i = 1, \ldots, k$, then for any $j = 1, \ldots, k$:

$$\mathbb{P}(B_j|A) = \frac{\mathbb{P}(A|B_j)\mathbb{P}(B_j)}{\sum_{i=1}^{k} \mathbb{P}(A|B_i)\mathbb{P}(B_i)}.$$

Proof:

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

where the first equality is by definition, and the second equality is by the multiplicative and total probability laws. $\square$

Corollary: If 0<P(B)<1, then

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

Proof: Take partition $\{B, \overline{B}\}$ in Bayes' rule. $\square$

Example

A firm rents cars from 3 agencies: 60% from agency 1, 30% from agency 2, and 10% from agency 3.
Suppose 9% of cars from agency 1 need a tuneup, 20% from agency 2 need a tuneup, and 6% from agency 3 need a tuneup.

If the rental car delivered to an employee needs a tuneup, what is the chance that it came from agency 2?

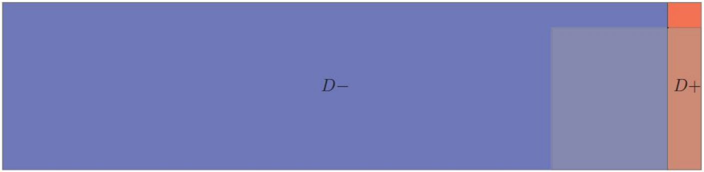
The Base Rate Fallacy

Example.
Consider a routine screening test for a disease. Suppose the frequency of the disease in the population (base rate) is 0.5%. The test is highly accurate with a 5% false positive rate and a 10% false negative rate.

You take the test and it comes back positive. What is the probability that you have the disease?

The Base Rate Fallacy

	D^+	D^-	total
T^+	45	498	543
T^-	5	9,452	9,457
total	50	9,950	10,000



The large blue area represents all the healthy people. The much smaller red area represents the sick people. The shaded rectangle represents the people who test positive.

Quiz

Monty Hall problem

Let's Make a Deal

1. One door hides a car, two hide goats.
2. Contestant chooses a door.
3. Monty (who knows where the car is) opens a different door with a goat.
4. Contestant can switch doors or keep her original choice.

What is the best strategy for winning a car?

(A) Switch

(B) Don't switch

(C) It doesn't matter

Quiz

Start with an urn with 5 red and 3 blue balls in it. Draw one ball. Put that ball back in the urn along with another ball of the same color. Now draw another ball from the urn.

Suppose the second ball is red. What is the probability the first ball was blue?
(A) 3/8
(B) 1/2
(C) 5/14
(D) 1/3
(E) 15/41