3 ed. Position

# Chapter 1
# Experiments, Models, and Probabilities

Chien-Kang Huang (黃乾綱)
臺大工科海洋系

# Before the Class

- Read the “A Message to Students from the Authors” in Preface.

# Outline

# Getting Started with Probability

- The probability of an event is a number between **0** and **1**.
  - The proportion of times we expect a certain thing to happen.
- Two interpretation of probability:
  - A physical property
  - Related to the prior knowledge
- Difference between probability theory and physics view point.
  - Probability theory – describe phenomena that cannot be predicted with certainty.
  - Physics – do the same thing in the same way, the result will always be the same.
- Probability: While each outcome may be **unpredictable**, there are **consistent patterns** to be observed when we repeat the procedure a large number of times.

# Three purposes served by this book

- It introduces students to the **logic** of probability theory
- It helps students develop **intuition** into how the theory applies to practical situations.
- It teaches students how to apply probability theory to **solving** engineering problems.

# Logic of the subject

- **Definitions**:
  - Establish the logic of probability theory
- **Axioms**:
  - facts that we accept without proof.
- **Theorems**:
  - Consequences that follow logically from definitions and axioms
  - Each theorem has a proof.
- There are only **3 axioms** of probability theory. These three axioms are the foundation on which the entire subject rests.
  - Dozens of Definitions → 3 Axioms → Dozens of Theorems

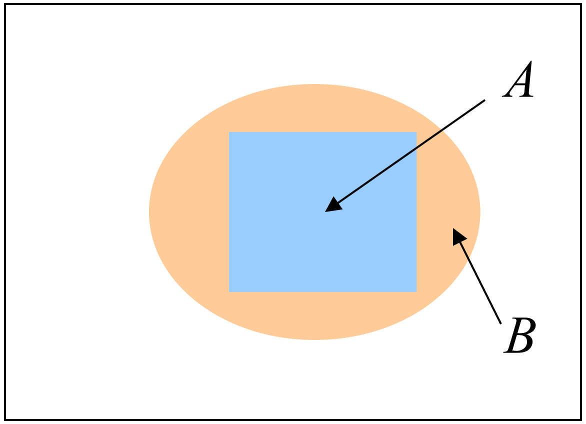
# 1.1 Set Theory

# Terminologies and Symbols

- Terms:
  - set, element, subset, universal set, null set ($\phi$), union, intersection, complement, differences, mutually exclusive, collectively exhaustive.
- Symbols:

| Symbols | Usage |
|---|---|
| $\in$ | $x \in A$ |
| $\notin$ | $c \notin A$ |
| $\{\}$ | {all the students in this room} |
| $\vert$ | $C = \{x^2 \mid x = 1, 2, 3, 4, 5\}$ |
| $\ldots$ | $D = \{x^2 \mid x = 1, 2, 3, 4, 5, \ldots\}$ |
| $\subset$ | $C \subset D$ |
| $=$ | $A = B$ iff $B \subset A$ and $A \subset B$ |
| $\phi$ | For any set $A$, $\phi \subset A$ |

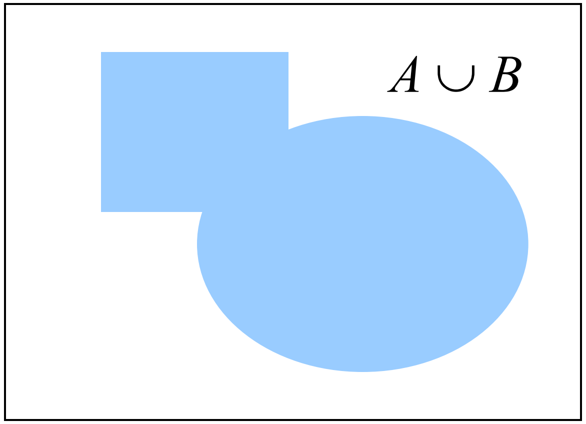
# Venn Diagram



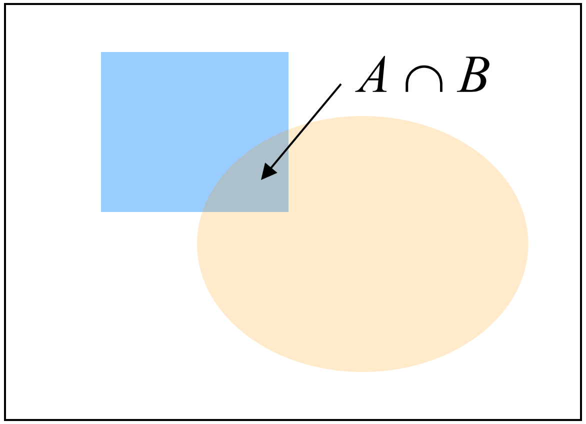

$A \subset B$

- if and only if, **iff**:
  - Symbols $\leftrightarrow, \Leftrightarrow, \equiv$
  - http://en.wikipedia.org/wiki/If_and_only_if

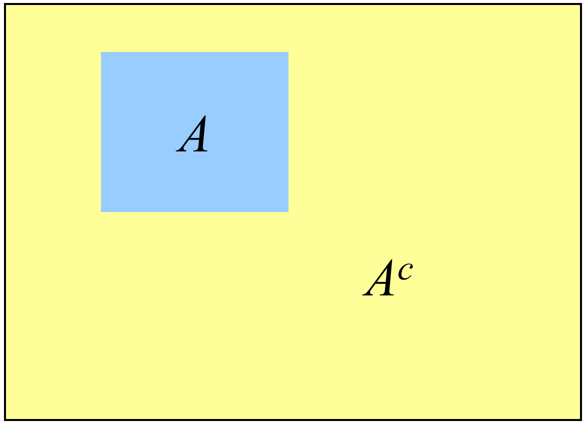
# Venn Diagram




union




intersection




complement

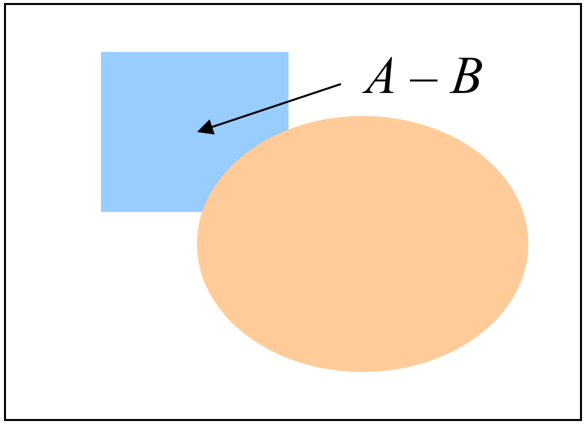
$$x \in A \cup B$$
$$\Updownarrow$$
$$x \in A \text{ or } x \in B$$

$$x \in A \cap B$$
$$\Updownarrow$$
$$x \in A \text{ and } x \in B$$

$$x \in A^c \Leftrightarrow x \notin A$$

Three operations

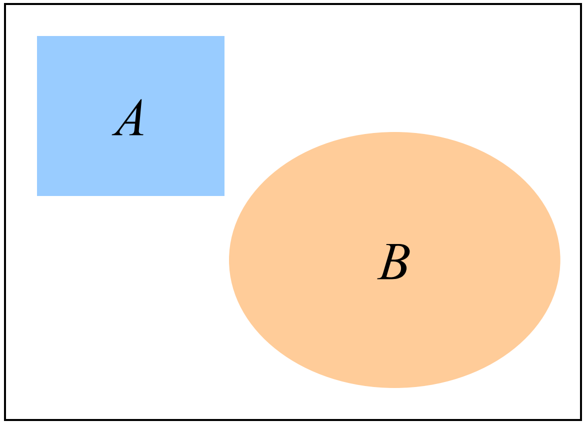
# Venn Diagram




difference

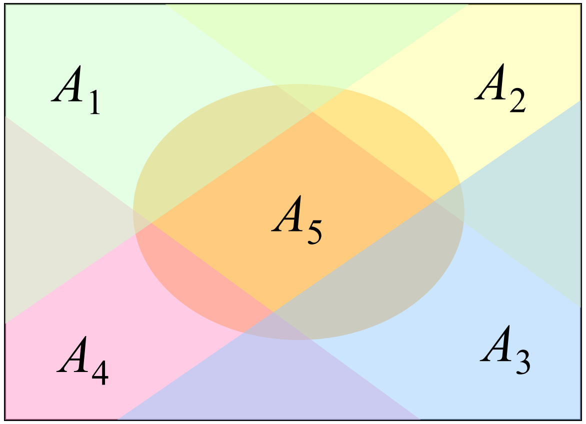
$$x \in A - B$$
$$\Updownarrow$$
$$x \in A \text{ and } x \notin B$$

$$A - B = A \cap B^C$$
$$B^C = S - B$$




($n$ sets) mutually exclusive,
(2 sets) disjoint

$$\Updownarrow$$
$$A_i \cap A_j = \phi, \; i \neq j.$$




collectively exhaustive

$$\Updownarrow$$
$$A_1 \cup A_2 \cup \cdots \cup A_n = S$$

# Shorthands

- Shorthand for unions and intersections of $n$ sets:

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \cdots \cup A_n$$

$$\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \cdots \cap A_n$$

# Theorem 1.1

De Morgan's law relates all three basic operations:

$$(A \cup B)^c = A^c \cap B^c.$$

# Proof: Theorem 1.1

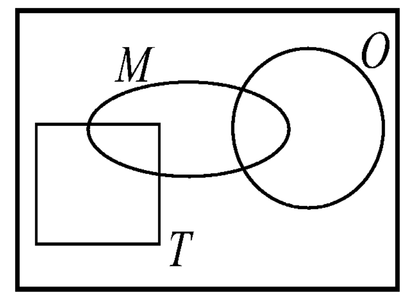
There are two parts to the proof:

- To show $(A \cup B)^c \subset A^c \cap B^c$, suppose $x \in (A \cup B)^c$. That implies $x \notin A \cup B$. Hence, $x \notin A$ and $x \notin B$, which together imply $x \in A^c$ and $x \in B^c$. That is, $x \in A^c \cap B^c$.

- To show $A^c \cap B^c \subset (A \cup B)^c$, suppose $x \in A^c \cap B^c$. In this case, $x \in A^c$ and $x \in B^c$. Equivalently, $x \notin A$ and $x \notin B$ so that $x \notin A \cup B$. Hence, $x \in (A \cup B)^c$.

# Quiz 1.1

A pizza at Gerlanda's is either regular ($R$) or Tuscan ($T$). In addition, each slice may have mushrooms ($M$) or onions ($O$) as described by the Venn diagram at right. For the sets specified below, shade the corresponding region of the Venn diagram.




(1) $R$

(2) $M \cup O$

(3) $M \cap O$

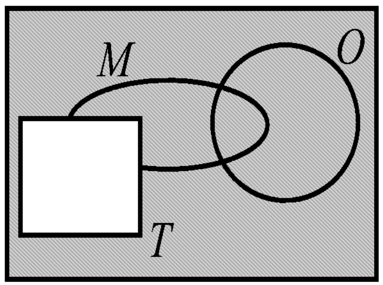
(4) $R \cup M$

(5) $R \cap M$

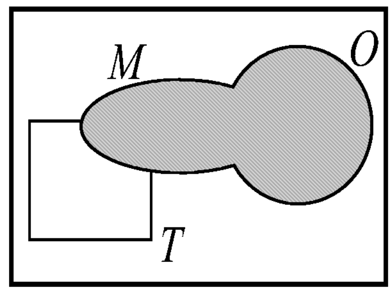
(6) $T^c - M$

# Quiz 1.1 Solution

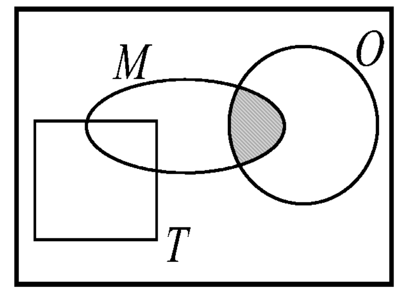
In the Venn diagrams for parts (a)-(g) below, the shaded area represents the indicated set.



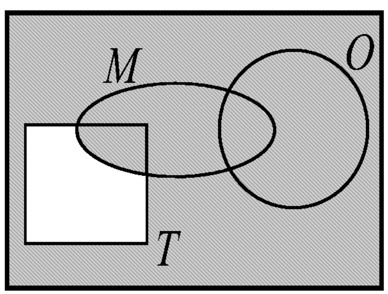

(1) $R = T^c$



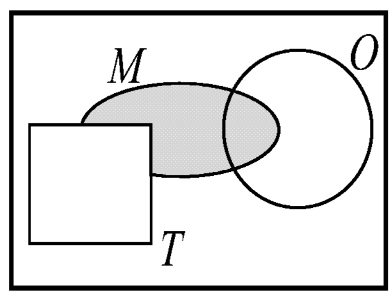

(2) $M \cup O$



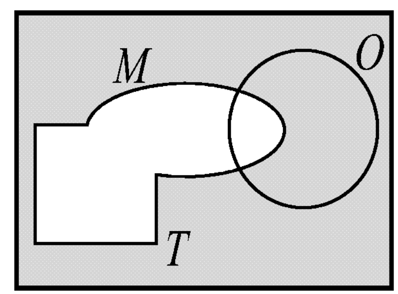

(3) $M \cap O$




(4) $R \cup M$




(4) $R \cap M$




(6) $T^c - M$

# 1.2 Applying Set Theory to Probability

# Terminologies

- Probability is a number that describe a set. The higher the number, the more probability there is.
- The basic model is a repeatable experiment. An **experiment** consists of a **procedure** and **observations**. There is uncertainty in what will be observed.
- For the most part, we will analyze **models** of actual physical experiments. We create models because real experiments generally are too complicated to analyze.
- We often will use the word experiment to refer to the model of an experiment.

## *Example 1.1*

An experiment consists of the following procedure, observation, and model:

- Procedure: Flip a coin and let it land on a table.
- Observation: Observe which side (head or tail) faces you after the coin lands.
- Model: Heads and tails are equally likely. The result of each flip is unrelated to the results of previous flips.

## *Example 1.2*

Flip a coin three times. Observe the sequence of heads and tails.

## *Example 1.3*

Flip a coin three times. Observe the number of heads.

## Definition 1.1 Outcome

*An* outcome *of an experiment is any possible observation of that experiment.*

## Definition 1.2 Sample Space

*The* sample space *of an experiment is the finest-grain, mutually exclusive, collectively exhaustive set of all possible outcomes.*

## Definition 1.3 Event

*An* event *is a set of outcomes of an experiment.*

## Definition 1.4 Event Space

*An* event space *is a collectively exhaustive, mutually exclusive set of events.*

| **Set Algebra** | **Probability** |
|---|---|
| Set | Event |
| Universal set | Sample space |
| Element | Outcome |

# *Example 1.4*

- The sample space in Example 1.1 is $S = \{h, t\}$ where $h$ is the outcome "observe head," and $t$ is the outcome "observe tail."

- The sample space in Example 1.2 is

$$S = \{hhh, hht, hth, htt, thh, tht, tth, ttt\}$$

- The sample space in Example 1.3 is $S = \{0, 1, 2, 3\}$.

# *Example 1.5*

Manufacture an integrated circuit and test it to determine whether it meets quality objectives. The possible outcomes are "accepted" $(a)$ and "rejected" $(r)$. The sample space is $S = \{a, r\}$.

# *Example 1.6*

Suppose we roll a six-sided die and observe the number of dots on the side facing upwards. We can label these outcomes $i = 1, \ldots, 6$ where $i$ denotes the outcome that $i$ dots appear on the up face. The sample space is $S = \{1, 2, \ldots, 6\}$. Each subset of $S$ is an event. Examples of events are

- The event $E_1 = \{\text{Roll 4 or higher}\} = \{4, 5, 6\}$.
- The event $E_2 = \{\text{The roll is even}\} = \{2, 4, 6\}$.
- $E_3 = \{\text{The roll is the square of an integer}\} = \{1, 4\}$.

# Example 1.7

Wait for someone to make a phone call and observe the duration of the call in minutes. An outcome $x$ is a nonnegative real number. The sample space is $S = \{x|x \geq 0\}$. The event "the phone call lasts longer than five minutes" is $\{x|x > 5\}$.

# *Example 1.8*

A short-circuit tester has a red light to indicate that there is a short circuit and a green light to indicate that there is no short circuit. Consider an experiment consisting of a sequence of three tests. In each test the observation is the color of the light that is on at the end of a test. An outcome of the experiment is a sequence of red ($r$) and green ($g$) lights. We can denote each outcome by a three-letter word such as $rgr$, the outcome that the first and third lights were red but the second light was green. We denote the event that light $n$ was red or green by $R_n$ or $G_n$. The event $R_2 = \{grg, grr, rrg, rrr\}$. We can also denote an outcome as an intersection of events $R_i$ and $G_j$. For example, the event $R_1G_2R_3$ is the set containing the single outcome $\{rgr\}$.

The set of events $\{G_2, R_2\}$ is both mutually exclusive and collectively exhaustive. However, $\{G_2, R_2\}$ is not a sample space for the experiment because the elements of the set do not completely describe the set of possible outcomes of the experiment. The set $\{G_2, R_2\}$ does not have the finest-grain property.

## Example 1.9 Problem

Flip four coins, a penny, a nickel, a dime, and a quarter. Examine the coins in order (penny, then nickel, then dime, then quarter) and observe whether each coin shows a head ($h$) or a tail ($t$). What is the sample space? How many elements are in the sample space?

## Example 1.9 Solution

The sample space consists of 16 four-letter words, with each letter either $h$ or $t$. For example, the outcome $tthh$ refers to the penny and the nickel showing tails and the dime and quarter showing heads. There are 16 members of the sample space.

## Example 1.10

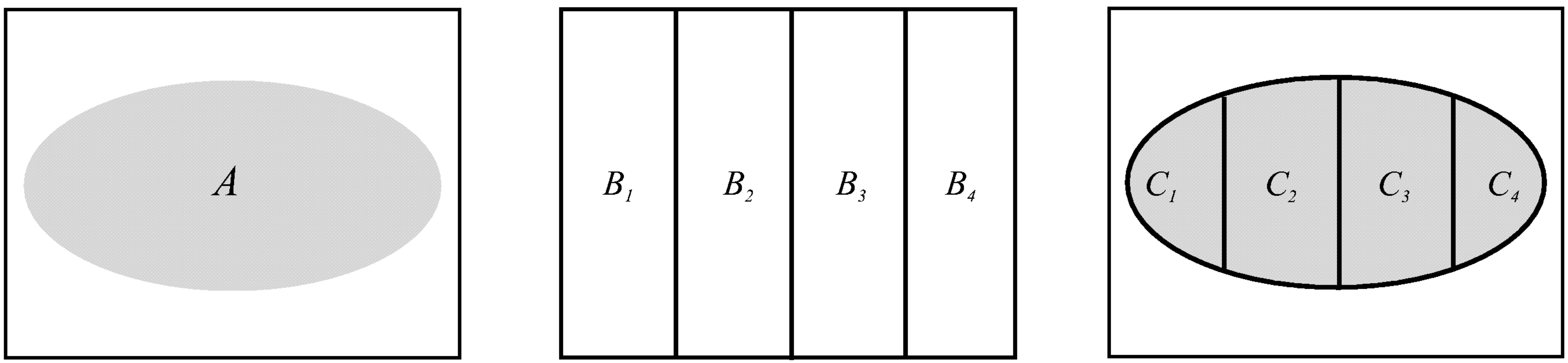
Continuing Example 1.9, let $B_i = \{\text{outcomes with } i \text{ heads}\}$. Each $B_i$ is an event containing one or more outcomes. For example, $B_1 = \{ttth, ttht, thtt, httt\}$ contains four outcomes. The set $B = \{B_0, B_1, B_2, B_3, B_4\}$ is an event space. Its members are mutually exclusive and collectively exhaustive. It is not a sample space because it lacks the finest-grain property. Learning that an experiment produces an event $B_1$ tells you that one coin came up heads, but it doesn't tell you which coin it was.

# Theorem 1.2

For an event space $B = \{B_1, B_2, \ldots\}$ and any event $A$ in the sample space, let $C_i = A \cap B_i$. For $i \neq j$, the events $C_i$ and $C_j$ are mutually exclusive and

$$A = C_1 \cup C_2 \cup \cdots .$$

# Figure 1.1




In this example of Theorem 1.2, the event space is $B = \{B_1, B_2, B_3, B_4\}$ and $C_i = A \cap B_i$ for $i = 1, \ldots, 4$. It should be apparent that $A = C_1 \cup C_2 \cup C_3 \cup C_4$.

# Example 1.11

In the coin-tossing experiment of Example 1.9, let $A$ equal the set of outcomes with less than three heads:

$$A = \{tttt, httt, thtt, ttht, ttth, hhtt, htht, htth, tthh, thth, thht\}.$$

From Example 1.10, let $B_i = \{\text{outcomes with } i \text{ heads}\}$. Since $\{B_0, \ldots, B_4\}$ is an event space, Theorem 1.2 states that

$$A = (A \cap B_0) \cup (A \cap B_1) \cup (A \cap B_2) \cup (A \cap B_3) \cup (A \cap B_4)$$

In this example, $B_i \subset A$, for $i = 0, 1, 2$. Therefore $A \cap B_i = B_i$ for $i = 0, 1, 2$. Also, for $i = 3$ and $i = 4$, $A \cap B_i = \phi$ so that $A = B_0 \cup B_1 \cup B_2$, a union of disjoint sets. In words, this example states that the event "less than three heads" is the union of events "zero heads," "one head," and "two heads."

## *Quiz 1.2*

Monitor three consecutive phone calls going through a telephone switching office. Classify each one as a voice call $(v)$ if someone is speaking, or a data call $(d)$ if the call is carrying a modem or fax signal. Your observation is a sequence of three letters (each letter is either $v$ or $d$). For example, two voice calls followed by one data call corresponds to $vvd$. Write the elements of the following sets:

(1) $A_1 = \{\text{first call is a voice call}\}$

(2) $B_1 = \{\text{first call is a data call}\}$

(3) $A_2 = \{\text{second call is a voice call}\}$

(4) $B_2 = \{\text{second call is a data call}\}$

(5) $A_3 = \{\text{all calls are the same}\}$

(6) $B_3 = \{\text{voice and data alternate}\}$

(7) $A_4 = \{\text{one or more voice calls}\}$

(8) $B_4 = \{\text{two or more data calls}\}$

For each pair of events $A_1$ and $B_1$, $A_2$ and $B_2$, and so on, identify whether the pair of events is either mutually exclusive or collectively exhaustive or both.

# *Quiz 1.2 Solution*

(1) $A_1 = \{vvv, vvd, vdv, vdd\}$

(2) $B_1 = \{dvv, dvd, ddv, ddd\}$

(3) $A_2 = \{vvv, vvd, dvv, dvd\}$

(4) $B_2 = \{vdv, vdd, ddv, ddd\}$

(5) $A_3 = \{vvv, ddd\}$

(6) $B_3 = \{vdv, dvd\}$

(7) $A_4 = \{vvv, vvd, vdv, dvv, vdd, dvd, ddv\}$

(8) $B_4 = \{ddd, ddv, dvd, vdd\}$

Recall that $A_i$ and $B_i$ are collectively exhaustive if $A_i \cup B_i = S$. Also, $A_i$ and $B_i$ are mutually exclusive if $A_i \cap B_i = \phi$. Since we have written down each pair $A_i$ and $B_i$ above, we can simply check for these properties.

The pair $A_1$ and $B_1$ are mutually exclusive and collectively exhaustive. The pair $A_2$ and $B_2$ are mutually exclusive and collectively exhaustive. The pair $A_3$ and $B_3$ are mutually exclusive but *not* collectively exhaustive. The pair $A_4$ and $B_4$ are not mutually exclusive since $dvd$ belongs to $A_4$ and $B_4$. However, $A_4$ and $B_4$ are collectively exhaustive.

# 1.3 Probability Axioms

Definition 1.4

# Definition 1.5 Axioms of Probability

*A probability measure $P[\cdot]$ is a function that maps events in the sample space to real numbers such that*

**Axiom 1** *For any event $A$, $P[A] \geq 0$.*

**Axiom 2** $P[S] = 1$.

**Axiom 3** *For any countable collection $A_1, A_2, \ldots$ of mutually exclusive events*

$$P[A_1 \cup A_2 \cup \cdots] = P[A_1] + P[A_2] + \cdots .$$

$S$: universal set, sample space
$s$: elements of $S$, outcomes
$A$: sets of elements, events

The probability of an event is the proportion of the time that event is observed in a large number of runs of the experiment. The is the *relative frequency* notion of probability.

Theorem 1.1, 1.2

# Theorem 1.3

For mutually exclusive events $A_1$ and $A_2$,

$$P[A_1 \cup A_2] = P[A_1] + P[A_2].$$

# Theorem 1.4

If $A = A_1 \cup A_2 \cup \cdots \cup A_m$ and $A_i \cap A_j = \phi$ for $i \neq j$, then

$$P[A] = \sum_{i=1}^{m} P[A_i].$$

# Theorem 1.5

The probability of an event $B = \{s_1, s_2, \ldots, s_m\}$ is the sum of the probabilities of the outcomes contained in the event:

$$P[B] = \sum_{i=1}^{m} P[\{s_i\}].$$

# Proof: Theorem 1.5

Each outcome $s_i$ is an event (a set) with the single element $s_i$. Since outcomes by definition are mutually exclusive, $B$ can be expressed as the union of $m$ disjoint sets:

$$B = \{s_1\} \cup \{s_2\} \cup \cdots \cup \{s_m\}$$

with $\{s_i\} \cap \{s_j\} = \phi$ for $i \neq j$. Applying Theorem 1.4 with $B_i = \{s_i\}$ yields

$$P[B] = \sum_{i=1}^{m} P[\{s_i\}].$$

# Comments on Notation

| Notation | Meaning |
|---|---|
| $P[\cdot]$ | The probability of an event |
| $P[A]$ | A function that transforms event $A$ to a number between $0$ and $1$. |
| $\{s_i\}$ | A set with the single element $s_i$ |
| $P[\{s_i\}]$<br>$P[s_i]$ | The probability of $\{s_i\}$. $P[\{s_i\}] = P[s_i]$ |
| $P[A \cap B]$<br>$P[A, B]$<br>$P[AB]$ | The probability of the intersection of two events.<br>$P[A \cap B] = P[A, B] = P[AB]$ |

# Example 1.12

Let $T_i$ denote the duration (in minutes) of the $i$th phone call you place today. The probability that your first phone call lasts less than five minutes and your second phone call lasts at least ten minutes is $P[T_1 < 5, T_2 \geq 10]$.

# Theorem 1.6

For an experiment with sample space $S = \{s_1, \ldots, s_n\}$ in which each outcome $s_i$ is equally likely,

$$P[s_i] = 1/n \qquad 1 \leq i \leq n.$$

# Proof: Theorem 1.6

Since all outcomes have equal probability, there exists $p$ such that $P[s_i] = p$ for $i = 1, \ldots, n$. Theorem 1.5 implies

$$P[S] = P[s_1] + \cdots + P[s_n] = np.$$

Since Axiom 2 says $P[S] = 1$, we must have $p = 1/n$.

# Example 1.13 Problem

As in Example 1.6, roll a six-sided die in which all faces are equally likely. What is the probability of each outcome? Find the probabilities of the events: "Roll 4 or higher," "Roll an even number," and "Roll the square of an integer."

# Example 1.13 Solution

The probability of each outcome is

$$P[i] = 1/6 \qquad i = 1, 2, \ldots, 6.$$

The probabilities of the three events are

- $P[\text{Roll 4 or higher}] = P[4] + P[5] + P[6] = 1/2$.
- $P[\text{Roll an even number}] = P[2] + P[4] + P[6] = 1/2$.
- $P[\text{Roll the square of an integer}] = P[1] + P[4] = 1/3$.

# *Quiz 1.3*

A student's test score $T$ is an integer between $0$ and $100$ corresponding to the experimental outcomes $s_0, \ldots, s_{100}$. A score of 90 to 100 is an $A$, 80 to 89 is a $B$, 70 to 79 is a $C$, 60 to 69 is a $D$, and below 60 is a failing grade of $F$. Given that all scores between 51 and 100 are equally likely and a score of 50 or less never occurs, find the following probabilities:

(1) $P[\{s_{79}\}]$

(2) $P[\{s_{100}\}]$

(3) $P[A]$

(4) $P[F]$

(5) $P[T \geq 80]$

(6) $P[T < 90]$

(7) $P[$a $C$ grade or better$]$

(8) $P[$student passes$]$

# *Quiz 1.3 Solution*

There are exactly 50 equally likely outcomes: $s_{51}$ through $s_{100}$. Each of these outcomes has probability $0.02$.

(1) $P[\{s_{79}\}] = 0.02$

(2) $P[\{s_{100}\}] = 0.02$

(3) $P[A] = P[\{s_{90}, \ldots, s_{100}\}] = 11 \times 0.02 = 0.22$

(4) $P[F] = P[\{s_{51}, \ldots, s_{59}\}] = 9 \times 0.02 = 0.18$

(5) $P[T \geq 80] = P[\{s_{80}, \ldots, s_{100}\}] = 21 \times 0.02 = 0.42$

(6) $P[T < 90] = P[\{s_{51}, s_{52}, \ldots, s_{89}\}] = 39 \times 0.02 = 0.78$

(7) $P[\text{a } C \text{ grade or better}] = P[\{s_{70}, \ldots, s_{100}\}] = 31 \times 0.02 = 0.62$

(8) $P[\text{student passes}] = P[\{s_{60}, \ldots, s_{100}\}] = 41 \times 0.02 = 0.82$

# 1.4 Some Consequences of the Axioms

# *Theorem 1.7*

The probability measure $P[\cdot]$ satisfies

(a) $P[\phi] = 0$.

(b) $P[A^c] = 1 - P[A]$.

(c) For any $A$ and $B$ (not necessarily disjoint),

$$P[A \cup B] = P[A] + P[B] - P[A \cap B].$$

(d) If $A \subset B$, then $P[A] \leq P[B]$.

Example 1.8

# Theorem 1.8

For any event $A$, and event space $\{B_1, B_2, \ldots, B_m\}$,

$$P[A] = \sum_{i=1}^{m} P[A \cap B_i].$$

# Proof: Theorem 1.8

The proof follows directly from Theorem 1.2 and Theorem 1.4. In this case, the disjoint sets are $C_i = \{A \cap B_i\}$ .

Example 1.15

# Example 1.14

A company has a model of telephone usage. It classifies all calls as either long ($l$), if they last more than three minutes, or brief ($b$). It also observes whether calls carry voice ($v$), data ($d$), or fax ($f$). This model implies an experiment in which the procedure is to monitor a call and the observation consists of the type of call, $v$, $d$, or $f$, and the length, $l$ or $b$. The sample space has six outcomes $S = \{lv, bv, ld, bd, lf, bf\}$. In this problem, each call is classifed in two ways: by length and by type. Using $L$ for the event that a call is long and $B$ for the event that a call is brief, $\{L, B\}$ is an event space. Similarly, the voice ($V$), data ($D$) and fax ($F$) classification is an event space $\{V, D, F\}$. The sample space can be represented by a table in which the rows and columns are labeled by events and the intersection of each row and column event contains a single outcome. The corresponding table entry is the probability of that outcome. In this case, the table is

| | $V$ | $D$ | $F$ |
|---|---|---|---|
| $L$ | 0.3 | 0.12 | 0.15 |
| $B$ | 0.2 | 0.08 | 0.15 |

For example, from the table we can read that the probability of a brief data call is $P[bd] = P[BD] = 0.08$. Note that $\{V, D, F\}$ is an event space corresponding to $\{B_1, B_2, B_3\}$ in Theorem 1.8. Thus we can apply Theorem 1.8 to find the probability of a long call:

$$P[L] = P[LV] + P[LD] + P[LF] = 0.57.$$



Briefing in next page.

# Example 1.14

- Model of telephone usage
  - Time: long ($l$) > 3 mins, brief ($b$)
  - Type: voice($v$), data($d$), fax($f$)
  - Sample space $S = \{lv, bv, ld, bd, lf, bf\}$
  - Event symbol: $L, B, V, D, F$

| | $V$ | $D$ | $F$ |
|---|---|---|---|
| $L$ | 0.3 | 0.12 | 0.15 |
| $B$ | 0.2 | 0.08 | 0.15 |

- The probability of a long call
  - $P[L] = P[LV] + P[LD] + P[LF] = 0.57$

# *Quiz 1.4*

Monitor a phone call. Classify the call as a voice call $(V)$ if someone is speaking, or a data call $(D)$ if the call is carrying a modem or fax signal. Classify the call as long $(L)$ if the call lasts for more than three minutes; otherwise classify the call as brief $(B)$. Based on data collected by the telephone company, we use the following probability model: $P[V] = 0.7$, $P[L] = 0.6$, $P[VL] = 0.35$. Find the following probabilities:

(1) $P[DL]$

(2) $P[D \cup L]$

(3) $P[VB]$

(4) $P[V \cup L]$

(5) $P[V \cup D]$

(6) $P[LB]$

# Quiz 1.4 Solution

We can describe this experiment by the event space consisting of the four possible events $VB$, $VL$, $DB$, and $DL$. We represent these events in the table:

| | $V$ | $D$ |
|---|---|---|
| $L$ | 0.35 | ? |
| $B$ | ? | ? |

In a roundabout way, the problem statement tells us how to fill in the table. In particular,

$$P[V] = 0.7 = P[VL] + P[VB]$$
$$P[L] = 0.6 = P[VL] + P[DL]$$

Since $P[VL] = 0.35$, we can conclude that $P[VB] = 0.35$ and that $P[DL] = 0.6 - 0.35 = 0.25$. This allows us to fill in two more table entries:

| | $V$ | $D$ |
|---|---|---|
| $L$ | 0.35 | 0.25 |
| $B$ | 0.35 | ? |

The remaining table entry is filled in by observing that the probabilities must sum to $1$.
[Continued]

# Quiz 1.4 Solution (continued)

This implies $P[DB] = 0.05$ and the complete table is

| | $V$ | $D$ |
|---|---|---|
| $L$ | 0.35 | 0.25 |
| $B$ | 0.35 | 0.05 |

Finding the various probabilities is now straightforward:

(1) $P[DL] = 0.25$

(2) $P[D \cup L] = P[VL] + P[DL] + P[DB] = 0.35 + 0.25 + 0.05 = 0.65.$

(3) $P[VB] = 0.35$

(4) $P[V \cup L] = P[V] + P[L] - P[VL] = 0.7 + 0.6 - 0.35 = 0.95$

(5) $P[V \cup D] = P[S] = 1$

(6) $P[LB] = P[LL^c] = 0$

# 1.5 Conditional Probability

# Prior Probability

- It is sometimes useful to interpret $P[A]$ as our knowledge of the occurrence of event $A$ before an experiment takes place.
- Thus $P[A]$ reflects our knowledge of the occurrence of $A$ **prior** to performing an experiment. Sometimes, we refer to $P[A]$ as the ***a priori* probability**, or the **prior probability**, of $A$.

# Conditional Probability

- In many practical situations, it is not possible to find out the precise outcome of an experiment. Rather than the outcome $s_i$, itself, we obtain information that the outcome is in the set $B$.
- Conditional probability describes out knowledge of $A$ when we know that $B$ has occurred but we still don't know the precise outcome.
- Notation: $P[A|B]$ : "the probability of $A$ given $B$"

Example 1.9

# Example 1.15

Consider an experiment that consists of testing two integrated circuits that come from the same silicon wafer, and observing in each case whether a circuit is accepted ($a$) or rejected ($r$). The sample space of the experiment is $S = \{rr, ra, ar, aa\}$. Let $B$ denote the event that the first chip tested is rejected. Mathematically, $B = \{rr, ra\}$. Similarly, let $A = \{rr, ar\}$ denote the event that the second circuit is a failure.

The circuits come from a high-quality production line. Therefore the prior probability $P[A]$ is very low. In advance, we are pretty certain that the second circuit will be accepted. However, some wafers become contaminated by dust, and these wafers have a high proportion of defective chips. Given the knowledge of event $B$ that the first chip was rejected, our knowledge of the quality of the second chip changes. With the event $B$ that the first chip is a reject, the probability $P[A|B]$ that the second chip will also be rejected is higher than the *a priori* probability $P[A]$ because of the likelihood that dust contaminated the entire wafer.



Briefing in next page.

# Example 1.15

- Testing two ICs come from the same silicon wafer.
- Observing: accepted ($a$), rejected ($r$)
- Sample space: $S = \{rr, ra, ar, aa\}$
- Event $B$ = {first chip tested is rejected} = $\{rr, ra\}$
- Event $A$ = {second chip tested is rejected} = $\{rr, ar\}$

- High quality → $P[A]$ is very low.
- $P[A|B]$ is higher than the a priori probability $P[A]$ because of the likelihood that dust contaminated the entire wafer.

# Definition 1.6 Conditional Probability

*The conditional probability of the event $A$ given the occurrence of the event $B$ is*

$$P[A|B] = \frac{P[AB]}{P[B]}.$$

# Theorem 1.9

A conditional probability measure $P[A|B]$ has the following properties that correspond to the axioms of probability.

Axiom 1: $P[A|B] \geq 0$.

Axiom 2: $P[B|B] = 1$.

Axiom 3: If $A = A_1 \cup A_2 \cup \cdots$ with $A_i \cap A_j = \phi$ for $i \neq j$, then

$$P[A|B] = P[A_1|B] + P[A_2|B] + \cdots$$

# Example 1.16 Problem

With respect to Example 1.15, consider the a priori probability model

$$P[rr] = 0.01, \quad P[ra] = 0.01, \quad P[ar] = 0.01, \quad P[aa] = 0.97.$$

Find the probability of $A =$ "second chip rejected" and $B =$ "first chip rejected." Also find the conditional probability that the second chip is a reject given that the first chip is a reject.

- $A$ = "second chip rejected", $P[A] =?$
- $B$ = "first chip rejected", $P[B] = ?$
- $P[AB] = ?$
- $P[A|B] = ?$

Go to Example 1.15

# Example 1.16 Solution

We saw in Example 1.15 that $A$ is the union of two disjoint events (outcomes) $rr$ and $ar$. Therefore, the a priori probability that the second chip is rejected is

$$P[A] = P[rr] + P[ar] = 0.02$$

This is also the a priori probability that the first chip is rejected:

$$P[B] = P[rr] + P[ra] = 0.02.$$

The conditional probability of the second chip being rejected given that the first chip is rejected is, by definition, the ratio of $P[AB]$ to $P[B]$, where, in this example,

$$P[AB] = P[\text{both rejected}] = P[rr] = 0.01$$

Thus

$$P[A|B] = \frac{P[AB]}{P[B]} = 0.01/0.02 = 0.5.$$

The information that the first chip is a reject drastically changes our state of knowledge about the second chip. We started with near certainty, $P[A] = 0.02$, that the second chip would not fail and ended with complete uncertainty about the quality of the second chip, $P[A|B] = 0.5$.

# Example 1.17 Problem

Shuffle a deck of cards and observe the bottom card. What is the conditional probability that the bottom card is the ace of clubs given that the bottom card is a black card?

- $A$ = “bottom card is the ace of clubs”
- $B$ = “bottom card is a black card”
- $P[A|B] =$

# *Example 1.17 Solution*

The sample space consists of the 52 cards that can appear on the bottom of the deck. Let $A$ denote the event that the bottom card is the ace of clubs. Since all cards are equally likely to be at the bottom, the probability that a particular card, such as the ace of clubs, is at the bottom is $P[A] = 1/52$. Let $B$ be the event that the bottom card is a black card. The event $B$ occurs if the bottom card is one of the 26 clubs or spades, so that $P[B] = 26/52$. Given $B$, the conditional probability of $A$ is

$$P[A|B] = \frac{P[AB]}{P[B]} = \frac{P[A]}{P[B]} = \frac{1/52}{26/52} = \frac{1}{26}.$$

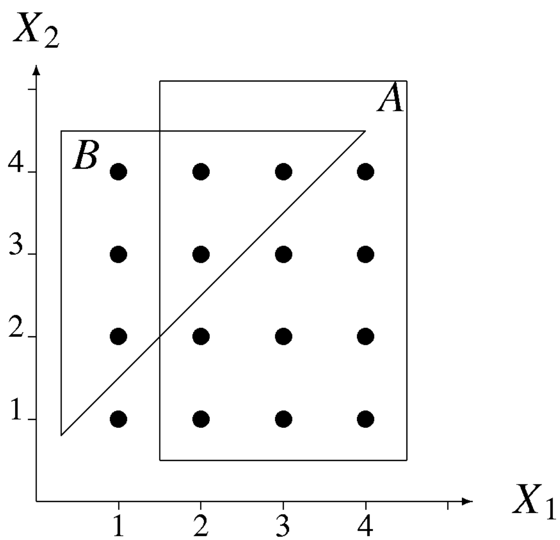
The key step was observing that $AB = A$, because if the bottom card is the ace of clubs, then the bottom card must be a black card. Mathematically, this is an example of the fact that $A \subset B$ implies that $AB = A$.

# Example 1.18 Problem

Roll two fair four-sided dice. Let $X_1$ and $X_2$ denote the number of dots that appear on die 1 and die 2, respectively. Let $A$ be the event $X_1 \geq 2$. What is $P[A]$? Let $B$ denote the event $X_2 > X_1$. What is $P[B]$? What is $P[A|B]$?

- $A = \{X_1 \geq 2\}$. $P[A] = ?$
- $B = \{X_2 > X_1\}$. $P[B] = ?$
- $P[A|B] = ?$

# Example 1.18 Solution




We begin by observing that the sample space has 16 elements corresponding to the four possible values of $X_1$ and the same four values of $X_2$. Since the dice are fair, the outcomes are equally likely, each with probability 1/16. We draw the sample space as a set of black circles in a two-dimensional diagram, in which the axes represent the events $X_1$ and $X_2$. Each outcome is a pair of values $(X_1, X_2)$. The rectangle represents $A$. It contains 12 outcomes, each with probability $1/16$.

To find $P[A]$, we add up the probabilities of outcomes in $A$, so $P[A] = 12/16 = 3/4$. The triangle represents $B$. It contains six outcomes. Therefore $P[B] = 6/16 = 3/8$. The event $AB$ has three outcomes, $(2,3), (2,4), (3,4)$, so $P[AB] = 3/16$. From the definition of conditional probability, we write

$$P[A|B] = \frac{P[AB]}{P[B]} = \frac{1}{2}.$$

We can also derive this fact from the diagram by restricting our attention to the six outcomes in $B$ (the conditioning event), and noting that three of the six outcomes in $B$ (one-half of the total) are also in $A$.

**Example 1.9**

## *Theorem 1.10 Law of Total Probability*

For an event space $\{B_1, B_2, \ldots, B_m\}$ with $P[B_i] > 0$ for all $i$,

$$P[A] = \sum_{i=1}^{m} P[A|B_i] P[B_i].$$

## *Proof: Theorem 1.10*

This follows from Theorem 1.8 and the identity $P[AB_i] = P[A|B_i]P[B_i]$, which is a direct consequence of the definition of conditional probability.

Example 1.16

# *Example 1.19 Problem*

A company has three machines $B_1$, $B_2$, and $B_3$ for making 1 k$\Omega$ resistors. It has been observed that 80% of resistors produced by $B_1$ are within 50 $\Omega$ of the nominal value. Machine $B_2$ produces 90% of resistors within 50 $\Omega$ of the nominal value. The percentage for machine $B_3$ is 60%. Each hour, machine $B_1$ produces 3000 resistors, $B_2$ produces 4000 resistors, and $B_3$ produces 3000 resistors. All of the resistors are mixed together at random in one bin and packed for shipment. What is the probability that the company ships a resistor that is within 50 $\Omega$ of the nominal value?

Briefing in next page.

# Example 1.19

- Three machines $B_1$, $B_2$, and $B_3$ for making 1k$\Omega$ resistors.
  - $B_1$: **80%** of resistors are within 50$\Omega$ of the nominal value.
    **3000** resistors per hour.
  - $B_2$: **90%** of resistors are within 50$\Omega$ of the nominal value.
    **4000** resistors per hour.
  - $B_3$: **60%** of resistors are within 50$\Omega$ of the nominal value.
    **3000** resistors per hour.
- What is the probability that the company ships a resistor that is within 50$\Omega$ of the nominal value?
- $A$ = "resistor is within 50$\Omega$ of the nominal value"
- $P[A] = ?$

# Example 1.19 Solution

Let $A = \{\text{resistor is within 50 } \Omega \text{ of the nominal value}\}$. Using the resistor accuracy information to formulate a probability model, we write

$$P[A|B_1] = 0.8, \quad P[A|B_2] = 0.9, \quad P[A|B_3] = 0.6$$

The production figures state that $3000 + 4000 + 3000 = 10{,}000$ resistors per hour are produced. The fraction from machine $B_1$ is $P[B_1] = 3000/10{,}000 = 0.3$. Similarly, $P[B_2] = 0.4$ and $P[B_3] = 0.3$. Now it is a simple matter to apply the law of total probability to find the accuracy probability for all resistors shipped by the company:

$$\begin{aligned} P[A] &= P[A|B_1]P[B_1] + P[A|B_2]P[B_2] + P[A|B_3]P[B_3] \\ &= (0.8)(0.3) + (0.9)(0.4) + (0.6)(0.3) = 0.78. \end{aligned}$$

For the whole factory, 78% of resistors are within 50 $\Omega$ of the nominal value.

Theorem 1.10

## Theorem 1.11 Bayes' theorem

$$P[B|A] = \frac{P[A|B]P[B]}{P[A]}.$$

## Proof: Theorem 1.11

$$P[B|A] = \frac{P[AB]}{P[A]} = \frac{P[A|B]P[B]}{P[A]}.$$

**Theorem 1.10 + 1.11**

$$P[B_i|A] = \frac{P[A|B_i]P[B_i]}{\sum_{i=1}^{m} P[A|B_i]P[B_i]}.$$

# *Example 1.20 Problem*

In Example 1.19 about a shipment of resistors from the factory, we learned that:

- The probability that a resistor is from machine $B_3$ is $P[B_3] = 0.3$.
- The probability that a resistor is *acceptable*, i.e., within 50 $\Omega$ of the nominal value, is $P[A] = 0.78$.
- Given that a resistor is from machine $B_3$, the conditional probability that it is acceptable is $P[A|B_3] = 0.6$.

What is the probability that an acceptable resistor comes from machine $B_3$?

# Example 1.20 Solution

Now we are given the event $A$ that a resistor is within 50 $\Omega$ of the nominal value, and we need to find $P[B_3|A]$. Using Bayes' theorem, we have

$$P[B_3|A] = \frac{P[A|B_3]P[B_3]}{P[A]}.$$

Since all of the quantities we need are given in the problem description, our answer is

$$P[B_3|A] = (0.6)(0.3)/(0.78) = 0.23.$$

Similarly we obtain $P[B_1|A] = 0.31$ and $P[B_2|A] = 0.46$. Of all resistors within 50 $\Omega$ of the nominal value, only 23% come from machine $B_3$ (even though this machine produces 30% of all resistors). Machine $B_1$ produces 31% of the resistors that meet the 50 $\Omega$ criterion and machine $B_2$ produces 46% of them.

# Quiz 1.5

Monitor three consecutive phone calls going through a telephone switching office. Classify each one as a voice call $(v)$ if someone is speaking, or a data call $(d)$ if the call is carrying a modem or fax signal. Your observation is a sequence of three letters (each one is either $v$ or $d$). For example, three voice calls corresponds to $vvv$. The outcomes $vvv$ and $ddd$ have probability $0.2$ whereas each of the other outcomes $vvd$, $vdv$, $vdd$, $dvv$, $dvd$, and $ddv$ has probability $0.1$. Count the number of voice calls $N_V$ in the three calls you have observed. Consider the four events $N_V = 0$, $N_V = 1$, $N_V = 2$, $N_V = 3$. Describe in words and also calculate the following probabilities:

(1) $P[N_V = 2]$

(2) $P[N_V \geq 1]$

(3) $P[\{vvd\}|N_V = 2]$

(4) $P[\{ddv\}|N_V = 2]$

(5) $P[N_V = 2|N_V \geq 1]$

(6) $P[N_V \geq 1|N_V = 2]$

# Quiz 1.5 Solution

(1) The probability of exactly two voice calls is

$$P[N_V = 2] = P[\{vvd, vdv, dvv\}] = 0.3$$

(2) The probability of at least one voice call is

$$\begin{aligned} P[N_V \geq 1] &= P[\{vdd, dvd, ddv, vvd, vdv, dvv, vvv\}] \\ &= 6(0.1) + 0.2 = 0.8 \end{aligned}$$

An easier way to get the same answer is to observe that

$$P[N_V \geq 1] = 1 - P[N_V < 1] = 1 - P[N_V = 0] = 1 - P[\{ddd\}] = 0.8$$

(3) The conditional probability of two voice calls followed by a data call given that there were two voice calls is

$$P[\{vvd\}|N_V = 2] = \frac{P[\{vvd\}, N_V = 2]}{P[N_V = 2]} = \frac{P[\{vvd\}]}{P[N_V = 2]} = \frac{0.1}{0.3} = \frac{1}{3}$$

(4) The conditional probability of two data calls followed by a voice call given there were two voice calls is

$$P[\{ddv\}|N_V = 2] = \frac{P[\{ddv\}, N_V = 2]}{P[N_V = 2]} = 0$$

The joint event of the outcome $ddv$ and exactly two voice calls has probability zero since there is only one voice call in the outcome $ddv$.

(5) The conditional probability of exactly two voice calls given at least one voice call is

$$P[N_V = 2|N_v \geq 1] = \frac{P[N_V = 2, N_V \geq 1]}{P[N_V \geq 1]} = \frac{P[N_V = 2]}{P[N_V \geq 1]} = \frac{0.3}{0.8} = \frac{3}{8}$$

(6) The conditional probability of at least one voice call given there were exactly two voice calls is

$$P[N_V \geq 1|N_V = 2] = \frac{P[N_V \geq 1, N_V = 2]}{P[N_V = 2]} = \frac{P[N_V = 2]}{P[N_V = 2]} = 1$$

Given that there were two voice calls, there must have been at least one voice call.

# 1.6 Independence

Definition 1.6

# *Definition 1.7 Two Independent Events*

*Events* $A$ *and* $B$ *are* independent *if and only if*

$$P[AB] = P[A]P[B].$$

- Equivalent definition of independent events

$$P[A|B] = P[A], \; P[B|A] = P[B]$$

- Independent and disjoint are not synonyms.
  - Disjoint events hare no outcomes in common, $P[AB] = 0$
  - Disjoint ➔ Axiom 3, **add** ⇔ **union**
  - Independent ➔ Definition 1.7, **multiply** ⇔ **intersection**

# Example 1.21 Problem

Suppose that for the three lights of Example 1.8, each outcome (a sequence of three lights, each either red or green) is equally likely. Are the events $R_2$ that the second light was red and $G_2$ that the second light was green independent? Are the events $R_1$ and $R_2$ independent?

- $S = \{rrr, rrg, rgr, rgg, grr, grg, ggr, ggg\}$
- $R_2 = \{rrr, rrg, grr, grg\}$, $G_2 = \{rgr, rgg, ggr, ggg\}$, $R_2 \cap G_2 = \phi$.
  $P[R_2G_2] = P[R_2]P[G_2]$ ?
- $R_2 = \{rrr, rrg, grr, grg\}$, $R_1 = \{rgg, rgr, rrg, rrr\}$, $R_2 \cap R_1 = \{rrg, rrr\}$
  $P[R_2R_1] = P[R_2]P[R_1]$ ?

# Example 1.21 Solution

Each element of the sample space

$$S = \{rrr, rrg, rgr, rgg, grr, grg, ggr, ggg\}$$

has probability $1/8$. Each of the events

$$R_2 = \{rrr, rrg, grr, grg\} \quad \text{and} \quad G_2 = \{rgr, rgg, ggr, ggg\}$$

contains four outcomes so $P[R_2] = P[G_2] = 4/8$. However, $R_2 \cap G_2 = \phi$ and $P[R_2G_2] = 0$. That is, $R_2$ and $G_2$ must be disjoint because the second light cannot be both red and green. Since $P[R_2G_2] \neq P[R_2]P[G_2]$, $R_2$ and $G_2$ are not independent. Learning whether or not the event $G_2$ (second light green) occurs drastically affects our knowledge of whether or not the event $R_2$ occurs. Each of the events $R_1 = \{rgg, rgr, rrg, rrr\}$ and $R_2 = \{rrg, rrr, grg, grr\}$ has four outcomes so $P[R_1] = P[R_2] = 4/8$. In this case, the intersection $R_1 \cap R_2 = \{rrg, rrr\}$ has probability $P[R_1R_2] = 2/8$. Since $P[R_1R_2] = P[R_1]P[R_2]$, events $R_1$ and $R_2$ are independent. Learning whether or not the event $R_2$ (second light red) occurs does not affect our knowledge of whether or not the event $R_1$ (first light red) occurs.

Example 1.19

# *Example 1.22 Problem*

Integrated circuits undergo two tests. A mechanical test determines whether pins have the correct spacing, and an electrical test checks the relationship of outputs to inputs. We *assume* that electrical failures and mechanical failures occur independently. Our information about circuit production tells us that mechanical failures occur with probability 0.05 and electrical failures occur with probability 0.2. What is the probability model of an experiment that consists of testing an integrated circuit and observing the results of the mechanical and electrical tests?

Briefing in next page.

# Example 1.22

- ICs undergo two test
- Mechanical test (detect the spacing between pins)
- Electrical test (check the input/output relationship)
- We assume that electrical failures and mechanical failures occur independently.
- Probability of mechanical failures = 0.05
- Probability of electrical failures = 0.02
- What is the probability model?

# Example 1.22 Solution

To build the probability model, we note that the sample space contains four outcomes:

$$S = \{(ma, ea), (ma, er), (mr, ea), (mr, er)\}$$

where $m$ denotes mechanical, $e$ denotes electrical, $a$ denotes accept, and $r$ denotes reject. Let $M$ and $E$ denote the events that the mechanical and electrical tests are acceptable. Our prior information tells us that $P[M^c] = 0.05$, and $P[E^c] = 0.2$. This implies $P[M] = 0.95$ and $P[E] = 0.8$. Using the independence assumption and Definition 1.7, we obtain the probabilities of the four outcomes in the sample space as

$$\begin{aligned}
P[(ma, ea)] &= P[ME] = P[M]P[E] = 0.95 \times 0.8 = 0.76, \\
P[(ma, er)] &= P[ME^c] = P[M]P[E^c] = 0.95 \times 0.2 = 0.19, \\
P[(mr, ea)] &= P[M^cE] = P[M^c]P[E] = 0.05 \times 0.8 = 0.04, \\
P[(mr, er)] &= P[M^cE^c] = P[M^c]P[E^c] = 0.05 \times 0.2 = 0.01.
\end{aligned}$$

Definition 1.7

# *Definition 1.8 3 Independent Events*

$A_1$, $A_2$, *and* $A_3$ *are* independent *if and only if*

*(a)* $A_1$ *and* $A_2$ *are independent,*

*(b)* $A_2$ *and* $A_3$ *are independent,*

*(c)* $A_1$ *and* $A_3$ *are independent,*

*(d)* $P[A_1 \cap A_2 \cap A_3] = P[A_1]P[A_2]P[A_3]$.

# Example 1.23 Problem

In an experiment with equiprobable outcomes, the event space is $S = \{1, 2, 3, 4\}$. $P[s] = 1/4$ for all $s \in S$. Are the events $A_1 = \{1, 3, 4\}$, $A_2 = \{2, 3, 4\}$, and $A_3 = \phi$ independent?

# Example 1.23 Solution

These three sets satisfy the final condition of Definition 1.8 because $A_1 \cap A_2 \cap A_3 = \phi$, and

$$P[A_1 \cap A_2 \cap A_3] = P[A_1]P[A_2]P[A_3] = 0.$$

However, $A_1$ and $A_2$ are not independent because, with all outcomes equiprobable,

$$P[A_1 \cap A_2] = P[\{3, 4\}] = 1/2 \neq P[A_1]P[A_2] = 3/4 \times 3/4.$$

Hence the three events are dependent.

## More than Two Independent

# Definition 1.9 Events

*If* $n \geq 3$, *the sets* $A_1, A_2, \ldots, A_n$ *are independent if and only if*

*(a) every set of* $n-1$ *sets taken from* $A_1, A_2, \ldots A_n$ *is independent,*

*(b)* $P[A_1 \cap A_2 \cap \cdots \cap A_n] = P[A_1]P[A_2]\cdots P[A_n]$.

# Quiz 1.6

Monitor two consecutive phone calls going through a telephone switching office. Classify each one as a voice call $(v)$ if someone is speaking, or a data call $(d)$ if the call is carrying a modem or fax signal. Your observation is a sequence of two letters (either $v$ or $d$). For example, two voice calls corresponds to $vv$. The two calls are independent and the probability that any one of them is a voice call is $0.8$. Denote the identity of call $i$ by $C_i$. If call $i$ is a voice call, then $C_i = v$; otherwise, $C_i = d$. Count the number of voice calls in the two calls you have observed. $N_V$ is the number of voice calls. Consider the three events $N_V = 0$, $N_V = 1$, $N_V = 2$. Determine whether the following pairs of events are independent:

(1) $\{N_V = 2\}$ and $\{N_V \geq 1\}$

(2) $\{N_V \geq 1\}$ and $\{C_1 = v\}$

(3) $\{C_2 = v\}$ and $\{C_1 = d\}$

(4) $\{C_2 = v\}$ and $\{N_V$ is even$\}$

# Quiz 1.6 Solution

In this experiment, there are four outcomes with probabilities

$$P[\{vv\}] = (0.8)^2 = 0.64 \qquad P[\{vd\}] = (0.8)(0.2) = 0.16$$
$$P[\{dv\}] = (0.2)(0.8) = 0.16 \qquad P[\{dd\}] = (0.2)^2 = 0.04$$

When checking the independence of any two events $A$ and $B$, it's wise to avoid intuition and simply check whether $P[AB] = P[A]P[B]$. Using the probabilities of the outcomes, we now can test for the independence of events.

[Continued]

# Quiz 1.6 Solution (continued)

(1) First, we calculate the probability of the joint event:

$$P[N_V = 2, N_V \geq 1] = P[N_V = 2] = P[\{vv\}] = 0.64$$

Next, we observe that

$$P[N_V \geq 1] = P[\{vd, dv, vv\}] = 0.96$$

Finally, we make the comparison

$$P[N_V = 2]P[N_V \geq 1] = (0.64)(0.96) \neq P[N_V = 2, N_V \geq 1]$$

which shows the two events are dependent.

(2) The probability of the joint event is

$$P[N_V \geq 1, C_1 = v] = P[\{vd, vv\}] = 0.80$$

From part (a), $P[N_V \geq 1] = 0.96$. Further, $P[C_1 = v] = 0.8$ so that

$$P[N_V \geq 1]P[C_1 = v] = (0.96)(0.8) = 0.768 \neq P[N_V \geq 1, C_1 = v]$$

Hence, the events are dependent.

(3) The problem statement that the calls were independent implies that the events the second call is a voice call, $\{C_2 = v\}$, and the first call is a data call, $\{C_1 = d\}$ are independent events. Just to be sure, we can do the calculations to check:

$$P[C_1 = d, C_2 = v] = P[\{dv\}] = 0.16$$

Since $P[C_1 = d]P[C_2 = v] = (0.2)(0.8) = 0.16$, we confirm that the events are independent. Note that this shouldn't be surprising since we used the information that the calls were independent in the problem statement to determine the probabilities of the outcomes.

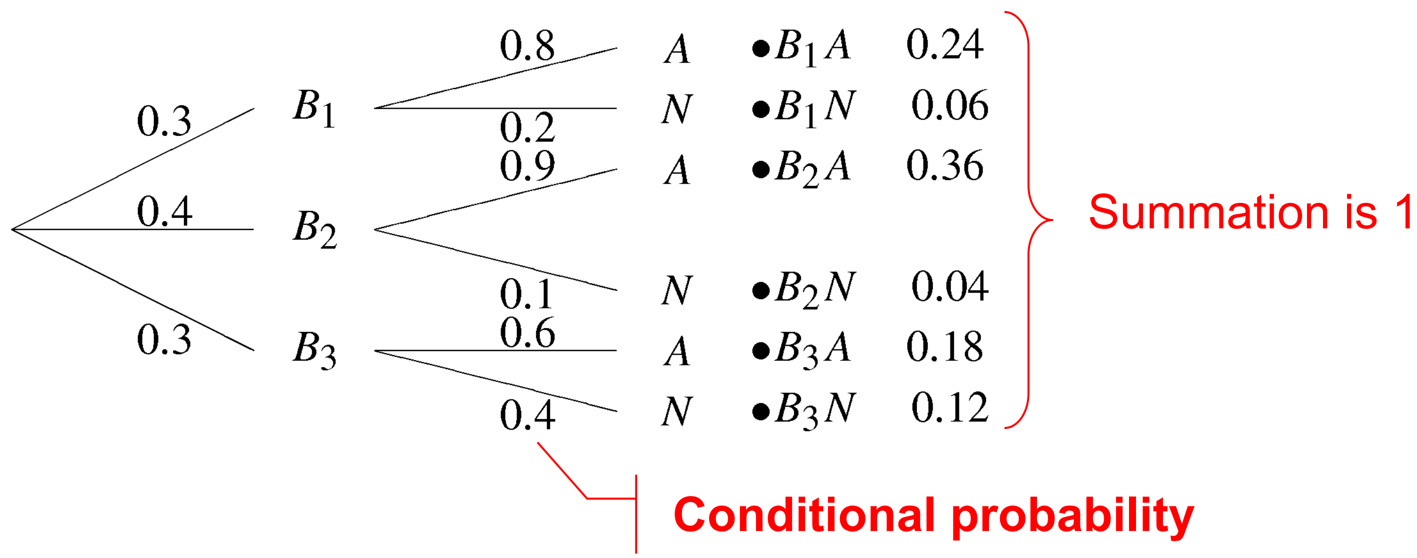
# 1.7 Sequential Experiments and Tree Diagrams

# Subexperiments and Tree Diagram

- Many experiments consist of a sequence of subexperiments.

## Figure 1.2

Theorem 1.10 + Theorem 1.9 (Axiom 3, Axiom 2)




The sequential tree for Example 1.24.

Example 2.1

# Example 1.24 Problem

For the resistors of Example 1.19, we have used $A$ to denote the event that a randomly chosen resistor is "within 50 $\Omega$ of the nominal value." This could mean "acceptable." We use the notation $N$ ("not acceptable") for the complement of $A$. The experiment of testing a resistor can be viewed as a two-step procedure. First we identify which machine ($B_1$, $B_2$, or $B_3$) produced the resistor. Second, we find out if the resistor is acceptable. Sketch a sequential tree for this experiment. What is the probability of choosing a resistor from machine $B_2$ that is not acceptable?

# Example 1.24 Solution

This two-step procedure corresponds to the tree shown in Figure 1.2. To use the tree to find the probability of the event $B_2N$, a nonacceptable resistor from machine $B_2$, we start at the left and find that the probability of reaching $B_2$ is $P[B_2] = 0.4$. We then move to the right to $B_2N$ and multiply $P[B_2]$ by $P[N|B_2] = 0.1$ to obtain $P[B_2N] = (0.4)(0.1) = 0.04$.

# Example 1.24

- For the resistors of Example 1.19
- $A$: a randomly chosen resistor is "within 50$\Omega$ of the nominal value."
- Two-step procedure for testing a resistor.
  - First, identify which machine ($B_1$, $B_2$, or $B_3$) produced the resistor.
  - Second, we find out if the resistor is acceptable.
- Sketch a sequential tree.
- What's the probability of choosing a resistor from machine $B_2$ that is not acceptable.

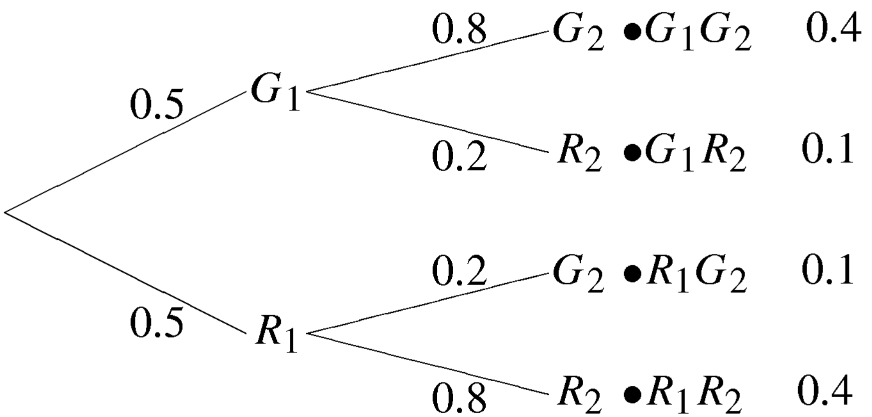
Example 2.2

# Example 1.25 Problem

Suppose traffic engineers have coordinated the timing of two traffic lights to encourage a run of green lights. In particular, the timing was designed so that with probability $0.8$ a driver will find the second light to have the same color as the first. Assuming the first light is equally likely to be red or green, what is the probability $P[G_2]$ that the second light is green? Also, what is $P[W]$, the probability that you wait for at least one light? Lastly, what is $P[G_1|R_2]$, the conditional probability of a green first light given a red second light?

# Example 1.25 Solution

In the case of the two-light experiment, the complete tree is




[Continued]

# Example 1.25 Solution (continued)

The probability that the second light is green is

$$P[G_2] = P[G_1G_2] + P[R_1G_2] = 0.4 + 0.1 = 0.5.$$

The event $W$ that you wait for at least one light is

$$W = \{R_1G_2 \cup G_1R_2 \cup R_1R_2\}.$$

The probability that you wait for at least one light is

$$P[W] = P[R_1G_2] + P[G_1R_2] + P[R_1R_2] = 0.1 + 0.1 + 0.4 = 0.6.$$

To find $P[G_1|R_2]$, we need $P[R_2] = 1 - P[G_2] = 0.5$. Since $P[G_1R_2] = 0.1$, the conditional probability that you have a green first light given a red second light is

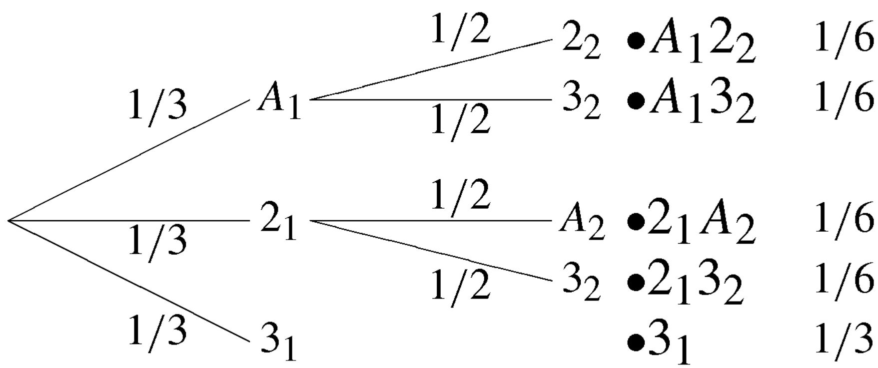
$$P[G_1|R_2] = \frac{P[G_1R_2]}{P[R_2]} = \frac{0.1}{0.5} = 0.2.$$

# Example 1.26 Problem

Consider the game of Three. You shuffle a deck of three cards: ace, 2, 3. With the ace worth 1 point, you draw cards until your total is 3 or more. You win if your total is 3. What is $P[W]$, the probability that you win?

# Example 1.26 Solution

Let $C_i$ denote the event that card $C$ is the $i$th card drawn. For example, $3_2$ is the event that the $3$ was the second card drawn. The tree is




You win if $A_1 2_2$, $2_1 A_2$, or $3_1$ occurs. Hence, the probability that you win is

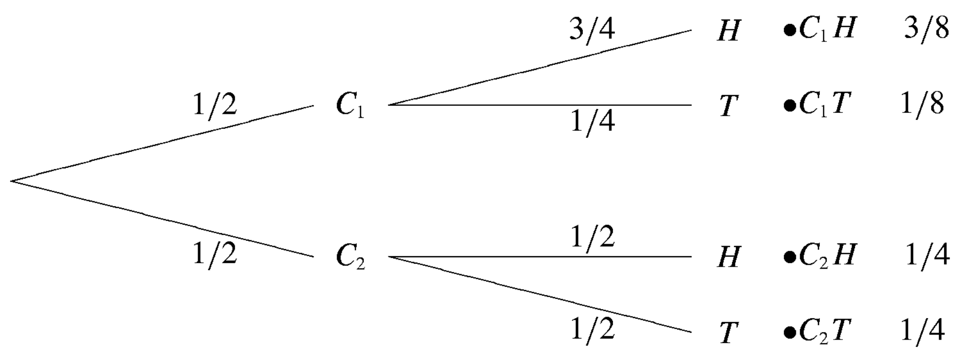
$$\begin{aligned} P[W] &= P[A_1 2_2] + P[2_1 A_2] + P[3_1] \\ &= (\frac{1}{3})(\frac{1}{2}) + (\frac{1}{3})(\frac{1}{2}) + \frac{1}{3} = \frac{2}{3}. \end{aligned}$$

# Example 1.27 Problem

Suppose you have two coins, one biased, one fair, but you don't know which coin is which. Coin 1 is biased. It comes up heads with probability $3/4$, while coin 2 will flip heads with probability $1/2$. Suppose you pick a coin at random and flip it. Let $C_i$ denote the event that coin $i$ is picked. Let $H$ and $T$ denote the possible outcomes of the flip. Given that the outcome of the flip is a head, what is $P[C_1|H]$, the probability that you picked the biased coin? Given that the outcome is a tail, what is the probability $P[C_1|T]$ that you picked the biased coin?

# Example 1.27 Solution

First, we construct the sample tree.




To find the conditional probabilities, we see

$$P[C_1|H] = \frac{P[C_1H]}{P[H]} = \frac{P[C_1H]}{P[C_1H]+P[C_2H]} = \frac{3/8}{3/8+1/4} = \frac{3}{5}.$$

Similarly,

$$P[C_1|T] = \frac{P[C_1T]}{P[T]} = \frac{P[C_1T]}{P[C_1T]+P[C_2T]} = \frac{1/8}{1/8+1/4} = \frac{1}{3}.$$

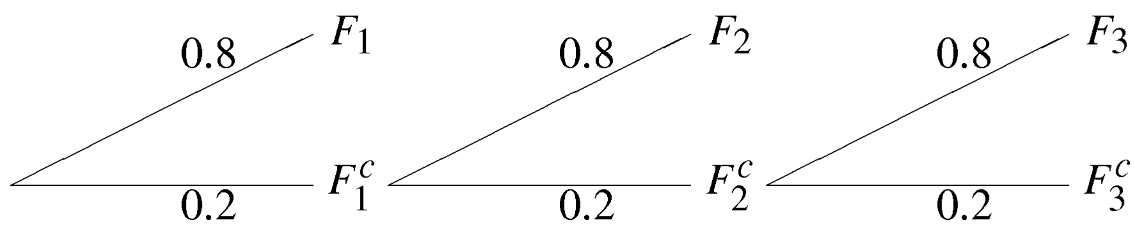
As we would expect, we are more likely to have chosen coin 1 when the first flip is heads, but we are more likely to have chosen coin 2 when the first flip is tails.

# Quiz 1.7

In a cellular phone system, a mobile phone must be paged to receive a phone call. However, paging attempts don't always succeed because the mobile phone may not receive the paging signal clearly. Consequently, the system will page a phone up to three times before giving up. If a single paging attempt succeeds with probability $0.8$, sketch a probability tree for this experiment and find the probability $P[F]$ that the phone is found.

# Quiz 1.7 Solution

Let $F_i$ denote the event that that the user is found on page $i$. The tree for the experiment is




The user is found unless all three paging attempts fail. Thus the probability the user is found is

$$P[F] = 1 - P[F_1^c F_2^c F_3^c] = 1 - (0.2)^3 = 0.992$$

# 1.8 Counting Method

Theorem 2.1

## Fundamental Principle of

## Definition 1.10 Counting

*If subexperiment $A$ has $n$ possible outcomes, and subexperiment $B$ has $k$ possible outcomes, then there are $nk$ possible outcomes when you perform both subexperiments.*

- Generally, if an experiment E has k subexperiments $E_1, \ldots, E_k$ where $E_i$ has $n_i$ outcomes, then E has

$$\prod_{i=1}^{k} n_i$$ outcomes.

Example 2.5

# Example 1.28

There are two subexperiments. The first subexperiment is "Flip a coin." It has two outcomes, $H$ and $T$. The second subexperiment is "Roll a die." It has six outcomes, $1, 2, \ldots, 6$. The experiment, "Flip a coin and roll a die," has $2 \times 6 = 12$ outcomes:

$$\begin{array}{cccccc} (H,1), & (H,2), & (H,3), & (H,4), & (H,5), & (H,6), \\ (T,1), & (T,2), & (T,3), & (T,4), & (T,5), & (T,6). \end{array}$$

## Example 1.29 Problem

Shuffle a deck and observe each card starting from the top. The outcome of the experiment is an ordered sequence of the 52 cards of the deck. How many possible outcomes are there?

## Example 1.29 Solution

The procedure consists of 52 subexperiments. In each one the observation is the identity of one card. The first subexperiment has 52 possible outcomes corresponding to the $52$ cards that could be drawn. After the first card is drawn, the second subexperiment has 51 possible outcomes corresponding to the 51 remaining cards. The total number of outcomes is

$$52 \times 51 \times \cdots \times 1 = 52!\,.$$

## Example 1.30 Problem

Shuffle the deck and choose three cards in order. How many outcomes are there?

## Example 1.30 Solution

In this experiment, there are 52 possible outcomes for the first card, 51 for the second card, and 50 for the third card. The total number of outcomes is $52 \times 51 \times 50$.

Theorem 2.2

# Permutation, Theorem 1.12

- An ordered sequence of $k$ **distinguishable objects** is called a ***k*-permutation**.
- Notation: $(n)_k = n(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!}$

## Theorem 1.12

The number of $k$-permutations of $n$ distinguishable objects is

$$(n)_k = n(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!}.$$

# Sampling without Replacement

- Replace choosing objects from a collection is also called **sampling**, and the chosen objects are know as a **sample**.
- Picking a subset of the collection of objects, each subset is called ***k*-combination**.
- Notation: $\binom{n}{k}$ read as "$n$ choose $k$"
- $k$-combination of $n$ objects + $k$-permutation of $k$ objects

$$(n)_k = \binom{n}{k} k!$$

# Theorem 1.13

The number of ways to choose $k$ objects out of $n$ distinguishable objects is

$$\binom{n}{k} = \frac{(n)_k}{k!} = \frac{n!}{k!(n-k)!}.$$

- **Binomial coefficient**: the coefficient of $x^k y^{n-k}$ in $(x+y)^n$
- Observation $\binom{n}{k} = \frac{(n)_k}{k!} = \frac{n!}{k!(n-k)!} = \binom{n}{n-k}$
- Constraint: $0 \le k \le n$. (Check the extended Definition 1.11)

Definition 2.1

# Definition 1.11 $n$ choose $k$

*For an integer* $n \geq 0$, *we define*

$$\binom{n}{k} = \begin{cases} \dfrac{n!}{k!(n-k)!} & k = 0, 1, \ldots, n, \\ 0 & \text{otherwise.} \end{cases}$$

Example 2.8

# Example 1.31

- The number of five-card poker hands is

$$\binom{52}{5} = \frac{52 \cdot 51 \cdot 50 \cdot 49 \cdot 48}{2 \cdot 3 \cdot 4 \cdot 5} = 2{,}598{,}960.$$

- The number of ways of picking 60 out of 120 students is $\binom{120}{60}$.

- The number of ways of choosing 5 starters for a basketball team with 11 players is $\binom{11}{5} = 462$.

- A baseball team has 15 field players and 10 pitchers. Each field player can take any of the 8 nonpitching positions. Therefore, the number of possible starting lineups is $N = \binom{10}{1}\binom{15}{8} = 64{,}350$ since you must choose 1 of the 10 pitchers and you must choose 8 out of the 15 field players. For each choice of starting lineup, the manager must submit to the umpire a batting order for the 9 starters. The number of possible batting orders is $N \times 9! = 23{,}351{,}328{,}000$ since there are $N$ ways to choose the 9 starters, and for each choice of 9 starters, there are $9! = 362{,}880$ possible batting orders.

# Example 1.32 Problem

To return to our original question of this section, suppose we draw seven cards. What is the probability of getting a hand without any queens?

# Example 1.32 Solution

There are $H = \binom{52}{7}$ possible hands. All $H$ hands have probability $1/H$. There are $H_{NQ} = \binom{48}{7}$ hands that have no queens since we must choose 7 cards from a deck of 48 cards that has no queens. Since all hands are equally likely, the probability of drawing no queens is $H_{NQ}/H = 0.5504$.

# Sampling with Replacement Theorem 1.14, 1.15

## Theorem 1.14

Given $m$ distinguishable objects, there are $m^n$ ways to choose with replacement an ordered sample of $n$ objects.

- Example 1.34 ~ 1.36

## Theorem 1.15

For $n$ repetitions of a subexperiment with sample space $S = \{s_0, \ldots, s_{m-1}\}$, there are $m^n$ possible observation sequences.

- Example 1.37 ~ 1.38

Example 2.11, 2.12

# Example 1.34

There are $2^{10} = 1024$ binary sequences of length $10$.

# Example 1.35

The letters $A$ through $Z$ can produce $26^4 = 456{,}976$ four-letter words.

# Example 1.36

A chip fabrication facility produces microprocessors. Each microprocessor is tested to determine whether it runs reliably at an acceptable clock speed. A subexperiment to test a microprocessor has sample space $S = \{0, 1\}$ to indicate whether the test was a failure $(0)$ or a success $(1)$. For test $i$, we record $x_i = 0$ or $x_i = 1$ to indicate the result. In testing four microprocessors, the observation sequence $x_1x_2x_3x_4$ is one of 16 possible outcomes:

$$\begin{array}{cccccccc} 0000, & 0001, & 0010, & 0011, & 0100, & 0101, & 0110, & 0111, \\ 1000, & 1001, & 1010, & 1011, & 1100, & 1101, & 1110, & 1111. \end{array}$$

# Example 1.37

A chip fabrication facility produces microprocessors. Each microprocessor is tested and assigned a grade $s \in S = \{s_0, \ldots, s_3\}$. A grade of $s_j$ indicates that the microprocessor will function reliably at a maximum clock rate of $s_j$ megahertz (MHz). In testing 10 microprocessors, we use $x_i$ to denote the grade of the $i$th microprocessor tested. Testing 10 microprocessors, for example, may produce an observation sequence

$$x_1 x_2 \cdots x_{10} = s_3 s_0 s_3 s_1 s_2 s_3 s_0 s_2 s_2 s_1.$$

The entire set of possible sequences contains $4^{10} = 1{,}048{,}576$ elements.

Example 2.14

# Example 1.38 Problem

For five subexperiments with sample space $S = \{0, 1\}$, how many observation sequences are there in which $0$ appears $n_0 = 2$ times and $1$ appears $n_1 = 3$ times?

# Example 1.38 Solution

The set of five-letter words with $0$ appearing twice and $1$ appearing three times is

$$\{00111, 01011, 01101, 01110, 10011, 10101, 10110, 11001, 11010, 11100\}.$$

There are exactly 10 such words.

# Theorem 1.16

The number of observation sequences for $n$ subexperiments with sample space $S = \{0, 1\}$ with $0$ appearing $n_0$ times and $1$ appearing $n_1 = n - n_0$ times is $\binom{n}{n_1}$.

# Theorem 1.17

For $n$ repetitions of a subexperiment with sample space $S = \{s_0, \ldots, s_{m-1}\}$, the number of length $n = n_0 + \cdots + n_{m-1}$ observation sequences with $s_i$ appearing $n_i$ times is

$$\binom{n}{n_0, \ldots, n_{m-1}} = \frac{n!}{n_0! n_1! \cdots n_{m-1}!}.$$

# Proof: Theorem 1.17

Let $M = \binom{n}{n_0,\ldots,n_{m-1}}$. Start with $n$ empty slots and perform the following sequence of subexperiments:

| Subexperiment | Procedure |
|---|---|
| $0$ | Label $n_0$ slots as $s_0$. |
| $1$ | Label $n_1$ slots as $s_1$. |
| $\vdots$ | $\vdots$ |
| $m-1$ | Label the remaining $n_{m-1}$ slots as $s_{m-1}$. |

There are $\binom{n}{n_0}$ ways to perform subexperiment $0$. After $n_0$ slots have been labeled, there are $\binom{n-n_0}{n_1}$ ways to perform subexperiment $1$. After subexperiment $j-1$, $n_0+\cdots+n_{j-1}$ slots have already been filled, leaving $\binom{n-(n_0+\cdots+n_{j-1})}{n_j}$ ways to perform subexperiment $j$. From the fundamental counting principle,

$$\begin{aligned} M &= \binom{n}{n_0}\binom{n-n_0}{n_1}\binom{n-n_0-n_1}{n_2}\cdots\binom{n-n_0-\cdots-n_{m-2}}{n_{m-1}} \\ &= \frac{n!}{(n-n_0)!n_0!}\frac{(n-n_0)!}{(n-n_0-n_1)!n_1!}\cdots\frac{(n-n_0-\cdots-n_{m-2})!}{(n-n_0-\cdots-n_{m-1})!n_{m-1}!}. \end{aligned}$$

Canceling the common factors, we obtain the formula of the theorem.

# Theorem 1.17

- For n repetitions of a subexperiment with sample space S = $\{s_0, \ldots, s_{m-1}\}$, the number of length n = $n_0 + \ldots + n_{m-1}$ observation sequences with $s_i$ appearing $n_i$ times is

$$\binom{n}{n_0, \cdots, n_{m-1}} = \frac{n!}{n_0!\, n_1! \cdots n_{m-1}!}$$

**Proof**

$$M = \binom{n}{n_0}\binom{n-n_0}{n_1}\binom{n-n_0-n_1}{n_2}\cdots\binom{n-n_0-\cdots-n_{m-2}}{n_{m-1}}$$

$$= \frac{n!}{n_0!\cancel{(n-n_0)!}}\,\frac{\cancel{(n-n_0)!}}{n_1!\cancel{(n-n_0-n_1)!}}\cdots\frac{\cancel{(n-n_0-\cdots-n_{m-2})!}}{n_{m-1}!\underbrace{(n-n_0-\cdots-n_{m-1})!}_{=\,0!\,=\,1}}$$

Definition 2.2

# Definition 1.12 Multinomial Coefficient

*For an integer* $n \geq 0$, *we define*

$$\binom{n}{n_0, \ldots, n_{m-1}} = \begin{cases} \dfrac{n!}{n_0! n_1! \cdots n_{m-1}!} & n_0 + \cdots + n_{m-1} = n; \\ & n_i \in \{0, 1, \ldots, n\}, i = 0, 1, \ldots, m-1, \\ 0 & \text{otherwise.} \end{cases}$$

Quiz 2.2

# Quiz 1.8

Consider a binary code with 4 bits (0 or 1) in each code word. An example of a code word is $0110$.

(1) How many different code words are there?

(2) How many code words have exactly two zeroes?

(3) How many code words begin with a zero?

(4) In a constant-ratio binary code, each code word has $N$ bits. In every word, $M$ of the $N$ bits are 1 and the other $N - M$ bits are 0. How many different code words are in the code with $N = 8$ and $M = 3$?

# Quiz 1.8 Solution

(1) We can view choosing each bit in the code word as a subexperiment. Each subexperiment has two possible outcomes: 0 and 1. Thus by the fundamental principle of counting, there are $2 \times 2 \times 2 \times 2 = 2^4 = 16$ possible code words.

(2) An experiment that can yield all possible code words with two zeroes is to choose which 2 bits (out of 4 bits) will be zero. The other two bits then must be ones. There are $\binom{4}{2} = 6$ ways to do this. Hence, there are six code words with exactly two zeroes. For this problem, it is also possible to simply enumerate the six code words:

$$1100, \quad 1010, \quad 1001, \quad 0101, \quad 0110, \quad 0011.$$

(3) When the first bit must be a zero, then the first subexperiment of choosing the first bit has only one outcome. For each of the next three bits, we have two choices. In this case, there are $1 \times 2 \times 2 \times 2 = 8$ ways of choosing a code word.

(4) For the constant ratio code, we can specify a code word by choosing $M$ of the bits to be ones. The other $N - M$ bits will be zeroes. The number of ways of choosing such a code word is $\binom{N}{M}$. For $N = 8$ and $M = 3$, there are $\binom{8}{3} = 56$ code words.

# 1.9 Independent Trials

# Theorem 1.18

The probability of $n_0$ failures and $n_1$ successes in $n = n_0 + n_1$ independent trials is

$$P[S_{n_0,n_1}] = \binom{n}{n_1}(1-p)^{n-n_1}p^{n_1} = \binom{n}{n_0}(1-p)^{n_0}p^{n-n_0}.$$

# Example 1.39 Problem

What is the probability $P[S_{2,3}]$ of two failures and three successes in five independent trials with success probability $p$.

# Example 1.39 Solution

To find $P[S_{2,3}]$, we observe that the outcomes with three successes in five trials are $11100$, $11010$, $11001$, $10110$, $10101$, $10011$, $01110$, $01101$, $01011$, and $00111$. We note that the probability of each outcome is a product of five probabilities, each related to one subexperiment. In outcomes with three successes, three of the probabilities are $p$ and the other two are $1-p$. Therefore each outcome with three successes has probability $(1-p)^2 p^3$.

From Theorem 1.16, we know that the number of such sequences is $\binom{5}{3}$. To find $P[S_{2,3}]$, we add up the probabilities associated with the 10 outcomes with 3 successes, yielding

$$P[S_{2,3}] = \binom{5}{3}(1-p)^2 p^3.$$

Example 2.17

# Example 1.40 Problem

In Example 1.19, we found that a randomly tested resistor was acceptable with probability $P[A] = 0.78$. If we randomly test 100 resistors, what is the probability of $T_i$, the event that $i$ resistors test acceptable?

# Example 1.40 Solution

Testing each resistor is an independent trial with a success occurring when a resistor is acceptable. Thus for $0 \leq i \leq 100$,

$$P[T_i] = \binom{100}{i} (0.78)^i (1 - 0.78)^{100-i}$$

We note that our intuition says that since 78% of the resistors are acceptable, then in testing 100 resistors, the number acceptable should be near 78. However, $P[T_{78}] \approx 0.096$, which is fairly small. This shows that although we might expect the number acceptable to be close to 78, that does not mean that the probability of exactly 78 acceptable is high.

# Example 1.41 Problem

To communicate one bit of information reliably, cellular phones transmit the same binary symbol five times. Thus the information "zero" is transmitted as 00000 and "one" is 11111. The receiver detects the correct information if three or more binary symbols are received correctly. What is the information error probability $P[E]$, if the binary symbol error probability is $q = 0.1$?

# Example 1.41 Solution

In this case, we have five trials corresponding to the five times the binary symbol is sent. On each trial, a success occurs when a binary symbol is received correctly. The probability of a success is $p = 1 - q = 0.9$. The error event $E$ occurs when the number of successes is strictly less than three:

$$\begin{aligned} P[E] &= P[S_{0,5}] + P[S_{1,4}] + P[S_{2,3}] \\ &= \binom{5}{0} q^5 + \binom{5}{1} p q^4 + \binom{5}{2} p^2 q^3 = 0.00856. \end{aligned}$$

By increasing the number of binary symbols per information bit from 1 to 5, the cellular phone reduces the probability of error by more than one order of magnitude, from 0.1 to 0.0081.

# Theorem 1.19

A subexperiment has sample space $S = \{s_0, \ldots, s_{m-1}\}$ with $P[s_i] = p_i$. For $n = n_0 + \cdots + n_{m-1}$ independent trials, the probability of $n_i$ occurences of $s_i$, $i = 0, 1, \ldots, m-1$, is

$$P[S_{n_0,\ldots,n_{m-1}}] = \binom{n}{n_0, \ldots, n_{m-1}} p_0^{n_0} \cdots p_{m-1}^{n_{m-1}}.$$

# Example 1.42

Each call arriving at a telephone switch is independently either a voice call with probability $7/10$, a fax call with probability $2/10$, or a modem call with probability $1/10$. Let $S_{v,f,m}$ denote the event that we observe $v$ voice calls, $f$ fax calls, and $m$ modem calls out of 100 observed calls. In this case,

$$P[S_{v,f,m}] = \binom{100}{v,\ f,\ m} (\frac{7}{10})^{v} (\frac{2}{10})^{f} (\frac{1}{10})^{m}$$

Keep in mind that by the extended definition of the multinomial coefficient, $P[S_{v,f,m}]$ is nonzero only if $v$, $f$, and $m$ are nonnegative integers such that $v + f + m = 100$.

## Example 1.43 Problem

Continuing with Example 1.37, suppose in testing a microprocessor that all four grades have probability $0.25$, independent of any other microprocessor. In testing $n = 100$ microprocessors, what is the probability of exactly $25$ microprocessors of each grade?

## Example 1.43 Solution

Let $S_{25,25,25,25}$ denote the probability of exactly 25 microprocessors of each grade. From Theorem 1.19,

$$P[S_{25,25,25,25}] = \binom{100}{25,\ 25,\ 25,\ 25}(0.25)^{100} = 0.0010.$$

# Quiz 1.9

Data packets containing $100$ bits are transmitted over a communication link. A transmitted bit is received in error (either a 0 sent is mistaken for a 1, or a 1 sent is mistaken for a 0) with probability $\epsilon = 0.01$, independent of the correctness of any other bit. The packet has been coded in such a way that if three or fewer bits are received in error, then those bits can be corrected. If more than three bits are received in error, then the packet is decoded with errors.

(1) Let $S_{k,100-k}$ denote the event that a received packet has $k$ bits in error and $100-k$ correctly decoded bits. What is $P[S_{k,100-k}]$ for $k = 0, 1, 2, 3$?

(2) Let $C$ denote the event that a packet is decoded correctly. What is $P[C]$?

# Quiz 1.9 Solution

(1) In this problem, $k$ bits received in error is the same as $k$ failures in $100$ trials. The failure probability is $\epsilon = 1 - p$ and the success probability is $1 - \epsilon = p$. That is, the probability of $k$ bits in error and $100 - k$ correctly received bits is

$$P[S_{k,100-k}] = \binom{100}{k} \epsilon^k (1-\epsilon)^{100-k}$$

For $\epsilon = 0.01$,

$$\begin{aligned}
P[S_{0,100}] &= (1-\epsilon)^{100} = (0.99)^{100} = 0.3660 \\
P[S_{1,99}] &= 100(0.01)(0.99)^{99} = 0.3700 \\
P[S_{2,98}] &= 4950(0.01)^2(0.99)^98 = 0.1849 \\
P[S_{3,97}] &= 161,700(0.01)^3(0.99)^{97} = 0.0610
\end{aligned}$$

(2) The probability a packet is decoded correctly is just

$$P[C] = P[S_{0,100}] + P[S_{1,99}] + P[S_{2,98}] + P[S_{3,97}] = 0.9819$$

# 1.10 Reliability Problems

# Reliability Problem

- The operation consists of $n$ components and each component succeeds with probability $p$, independent of any other component. Let $W_i$ denote the event that component $i$ succeeds.

- **Components in series**
  - The operation succeeds if **all** of its components succeed.

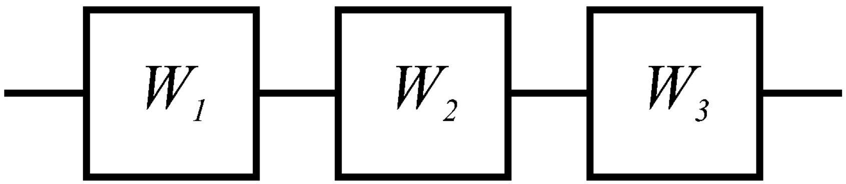
$$P[W] = P[W_1 W_2 \cdots W_n] = p \times p \times \cdots \times p = p^n$$

- **Components in parallel**
  - The operation succeeds if **any** of its components succeed.

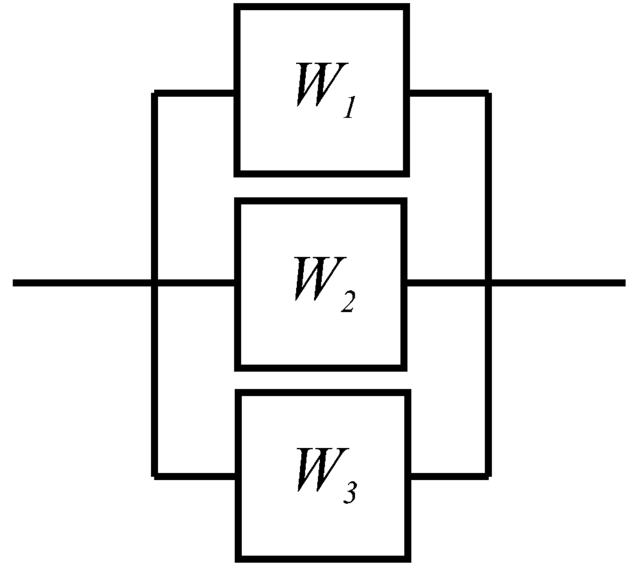
$$P[W^c] = P[W_1^c W_2^c \cdots W_n^c] = (1 - p)^n$$

$$P[W] = 1 - P[W^c] = 1 - (1 - p)^n$$

# Figure 1.3




*Components in Series*



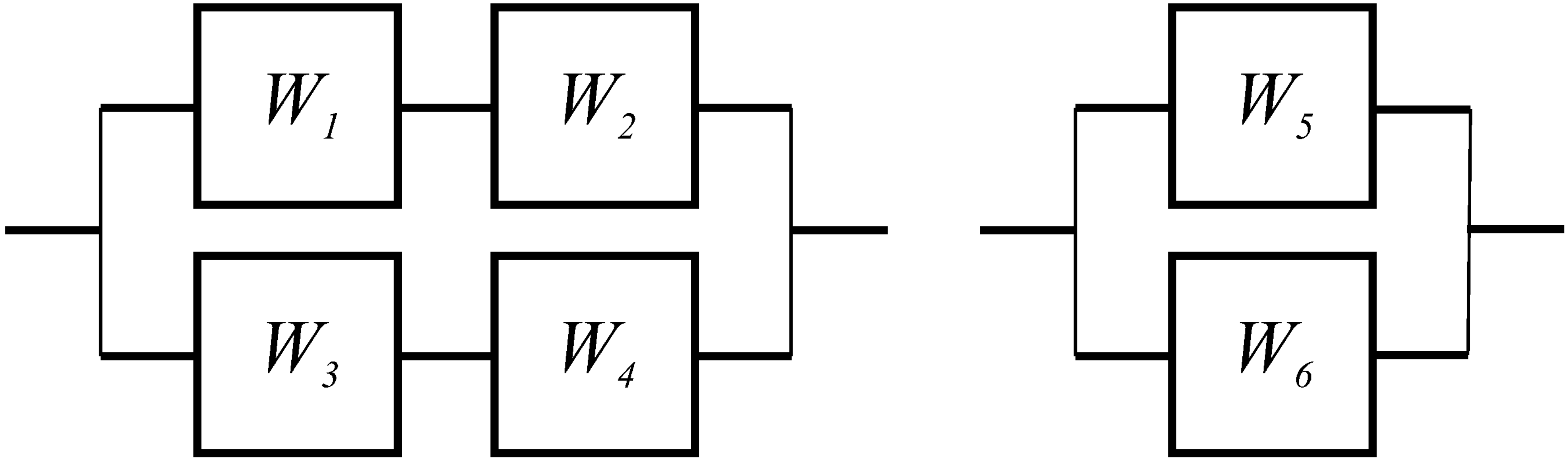

*Components in Parallel*

Serial and parallel devices.

$$P[W] = P[W_1 W_2 \cdots W_n]$$
$$= p \times p \times \cdots \times p = p^n$$

$$P[W^c] = P[W_1^c W_2^c \cdots W_n^c] = (1-p)^n$$
$$P[W] = 1 - P[W^c] = 1 - (1-p)^n$$

# Figure 1.4




The operation described in Example 1.44. On the left is the original operation. On the right is the equivalent operation with each pair of series components replaced with an equivalent component.

# Example 1.44 Problem

An operation consists of two redundant parts. The first part has two components in series ($W_1$ and $W_2$) and the second part has two components in series ($W_3$ and $W_4$). All components succeed with probability $p = 0.9$. Draw a diagram of the operation and calculate the probability that the operation succeeds.

# Example 1.44 Solution

A diagram of the operation is shown in Figure 1.4. We can create an equivalent component, $W_5$, with probability of success $p_5$ by observing that for the combination of $W_1$ and $W_2$,

$$P[W_5] = p_5 = P[W_1 W_2] = p^2 = 0.81.$$

Similarly, the combination of $W_3$ and $W_4$ in series produces an equivalent component, $W_6$, with probability of success $p_6 = p_5 = 0.81$. The entire operation then consists of $W_5$ and $W_6$ in parallel which is also shown in Figure 1.4. The success probability of the operation is

$$P[W] = 1 - (1 - p_5)^2 = 0.964$$

We could consider the combination of $W_5$ and $W_6$ to be an equivalent component $W_7$ with success probability $p_7 = 0.964$ and then analyze a more complex operation that contains $W_7$ as a component.

# Quiz 1.10

A memory module consists of nine chips. The device is designed with redundancy so that it works even if one of its chips is defective. Each chip contains $n$ transistors and functions properly if all of its transistors work. A transistor works with probability $p$ independent of any other transistor. What is the probability $P[C]$ that a chip works? What is the probability $P[M]$ that the memory module works?

# Quiz 1.10 Solution

Since the chip works only if all $n$ transistors work, the transistors in the chip are like devices in series. The probability that a chip works is $P[C] = p^n$.

The module works if either 8 chips work or 9 chips work. Let $C_k$ denote the event that exactly $k$ chips work. Since transistor failures are independent of each other, chip failures are also independent. Thus each $P[C_k]$ has the binomial probability

$$P[C_8] = \binom{9}{8}(P[C])^8(1 - P[C])^{9-8} = 9p^{8n}(1 - p^n),$$

$$P[C_9] = (P[C])^9 = p^{9n}.$$

The probability a memory module works is

$$P[M] = P[C_8] + P[C_9] = p^{8n}(9 - 8p^n)$$