STAT 509: Statistics for Engineers

Chapter 2: Probability

Dr. Dewei Wang
Associate Professor
Department of Statistics
University of South Carolina
deweiwang@stat.sc.edu

Chapter 2: Probability

Learning Objectives:

1. Understand and describe sample spaces and events
2. Interpret probabilities and calculate probabilities of events
3. Use permutations and combinations to count outcomes
4. Calculate the probabilities of joint events
5. Interpret and calculate conditional probabilities
6. Determine independence and use independence to calculate probabilities
7. Understand Bayes' theorem and when to use it

Random Experiment

An **experiment** is a procedure that is

- carried out under controlled conditions, and
- executed to discover an unknown result.

An experiment that results in different outcomes even when repeated in the same manner every time is a **random experiment**; e.g.,

- Flip a coin
- Toss a dice
- Measure the recycle time of a flash

How to describe the likelihood of observing a possible outcome from a random experiment? What is the probability of a "head" from a coin flipping?

The **set** of all possible outcomes of a random experiment is called the **sample space**, denoted by S.

- S is **discrete** if it consists of a finite or countable infinite set of outcomes.
- S is **continuous** if it contains an interval of real numbers.

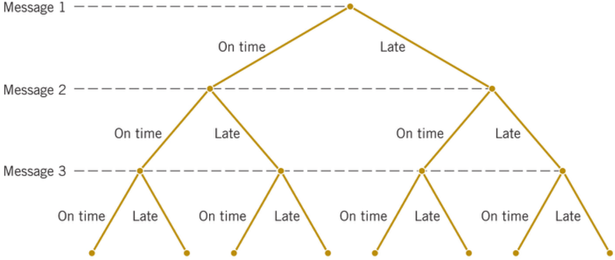
Examples:

1. Randomly select a camera and record the recycle time of a flash: $S = \mathbb{R}^+ = (0, \infty)$, all the positive real numbers, is continuous.
2. Suppose we know all the recycle times are between 1.5 and 5 seconds. Then $S = (1.5, 5)$ is continuous.
3. It is known that the recycle time has only three values(low, medium or high). Then $S = \{\text{low, medium, high}\}$ is discrete.
4. Does the camera conform to minimum recycle time specifications? $S = \{\text{yes, no}\}$ is discrete.

Tree diagram to list a discrete sample space

Messages are classified as on-time(o) or late(l). Classify the next 3 messages.

$$S = \{\text{ooo, ool, olo, oll, loo, lol, llo, lll}\}.$$



This only works for small sample spaces. Think we have 30 messages, the size of S is $2^{30} = 1,073,741,824$.

Counting Techniques

There are three special rules, or counting techniques, used to determine the number of outcomes in events:

1. Multiplication rule
2. Permutation rule
3. Combination rule

Each has its special purpose that must be applied properly – the right tool for the right job.

Multiplication Rule

Let an operation consists of k steps and there are

- n_1 ways of completing step 1,
- n_2 ways of completing step 2, ..., and
- n_k ways of completing step k.

Then, the total number of ways to perform this operation is

$$n_1 \cdot n_2 \cdots n_k.$$

Example: Web Site Design

In the design for a website, we can choose to use among: 4 colors, 3 fonts, and 3 positions for an image. How many designs are possible?
Answer via the multiplication rule: $4 \cdot 3 \cdot 3 = 36$.

Permutation Rule

A permutation is a unique sequence (**order matters**) of distinct items. For example, if $S = \{a, b, c\}$, there are $6 = 3 \times 2 \times 1$ permutations:

$$abc, acb, bac, bca, cab, cba.$$

How many different ways to permute n different items? Answer is

$$n! \text{ (factorial)} = n(n-1)(n-2)\cdots 2 \cdot 1.$$

by definition $0! = 1$.

Subset Permutations

How many different ways to permute r items from a set of n distinct items?

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

nPr(n,r)

Example

A printed circuit board has eight different locations in which a component can be placed. If four different components are to be placed on the board, how many designs are possible?

Answer: Order is important! Using the permutation formula with $n = 8$, $r = 4$:

$$P_4^8 = \frac{8!}{(8-4)!} = 8 \cdot 7 \cdot 6 \cdot 5 = 1680.$$

nPr(8,4)

Similar Item (not distinct) Permutations

Suppose the n items are not totally distinct. We have

- $n = n_1 + n_2 + \cdots + n_r$ items of which
- $n_1, n_2, \ldots, n_r$ are identical.

The number of permutations of these n items is

$$\frac{n!}{n_1! n_2! \cdots n_r!}$$

SimPerm(c(n1,n2,...,nr))

Example

In a hospital, an operating room needs to schedule 2 (identical) brain surgeries, 3 (identical) knee surgeries and 2 (identical) hip surgeries in a day. How many schedules are there?

$$\frac{(2+3+2)!}{2!3!2!} = 210.$$

SimPerm(c(2,3,2))

Combination Rule

A combination is a selection of r items from a set of n where **order does not matter**.

Example

If $S = \{a, b, c\}$, $n = 3$. Then

- If pick $r = 3$ out, we have 1 combination: abc (the same as acb, bca,...)
- If pick $r = 2$ out, we have 3 combinations: ab, bc, ac.

The number of permutations (where order matters) is always larger or equal to the number of combinations (where order does not matter).

The number of combinations of r times out of n is

$$C_r^n = \frac{n!}{r!(n-r)!}$$

nCr(n,r)

Example: Combination Rule

A bin of 50 parts contains 3 defectives and 47 non-defective parts. A sample of 6 parts is selected from the 50 **without replacement**. How many ways to get a sample of size 6 which contains 2 defective parts?

Answer:
Step 1: We need to sample 2 defectives out of the 3 defectives, which has $C_2^3 = 3$ different ways.
Step 2: To sample the remaing 4 non-defective parts out of the toal 47 ones, which has $C_4^{47} = 178,365$ different ways.

Thus, in total, there are $C_2^3 \times C_4^{47} = 3 \times 178,365 = 535,095$ different ways.
nCr(3,2)*nCr(47,4)

Events and Set Operations

An **event** (E) is a **subset** of the sample space of a random experiment.

Event combinations (set operations)

- The **Union** of two events, E_1 and E_2, consists of all outcomes that are contained in one event **or** the other, denoted as $E_1 \cup E_2$.
- The **Intersection** of two events E_1 and E_2, consists of all outcomes that are contained in one event **and** the other, denoted as $E_1 \cap E_2$.
- The **Complement** of an event E is the set of outcomes in the sample space that are **not** contained in the event, denoted as E^c.

Example: Discrete Events

Suppose that the recycle times of two cameras are recorded. Consider only whether or not the cameras conform to the manufacturing specifications. We abbreviate yes and no as y and n. The sample space is $S = \{yy, yn, ny, nn\}$. Let

- E_1 denote an event that at least one camera conforms to specifications, then $E_1 = \{yy, yn, ny\}$,
- E_2 an event that no camera conforms to specifications, then $E_2 = \{nn\}$,
- and E_3 an event that at least one camera does not conform, then $E_3 = \{yn, ny, nn\}$.

We have

- $E_1 \cup E_3 = S$
- $E_1 \cap E_3 = \{yn, ny\}$
- $E_1^c = \{nn\}$

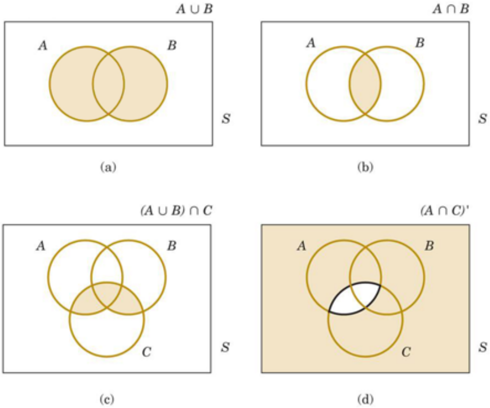
Example: Continuous Events

Measurements of the thickness of a part are modeled with the sample space: $S = (0, \infty)$. Let $E_1 = [10, 12)$ and $E_2 = (11, 15)$. Then

- $E_1 \cup E_2 = [10, 15)$
- $E_1 \cap E_2 = (11, 12)$
- $E_1^c = (0, 10) \cup [12, \infty)$
- $E_1^c \cap E_2 = [12, 15)$

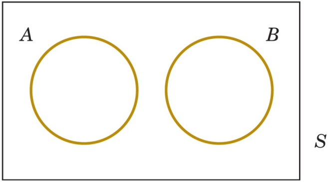
Venn Diagrams

Events A and B contain their respective outcomes. The shaded regions indicate the event relation of each diagram.



Mutually Exclusive Events

Events A and B are **mutually exclusive** because they share no common outcomes. The occurrence of one event precludes the occurrence of the other (not independent at all, strongly dependent). Symbolically, $A \cap B = \emptyset$ (the emptyset set).



Some laws of set operations

- Commutative law:
  $A \cap B = B \cap A$
  $A \cup B = B \cup A$.
- Distributive law:
  $A \cap (B \cup C) = (A \cup C) \cap (A \cup B)$
  $A \cup (B \cap C) = (A \cap C) \cup (A \cap B)$
- Associative law:
  $(A \cap B) \cap C = A \cap (B \cap C)$
  $(A \cup B) \cup C = A \cup (B \cup C)$
- Complement law: $(A^c)^c = A$
- De Morgan's law:
  $(A \cup B)^c = A^c \cap B^c$
  $(A \cap B)^c = A^c \cup B^c$

Probability

Probability is the likelihood or chance that a particular outcome or event from a random experiment will occur.

Denote by $P(E)$ the probability of event E will occur.

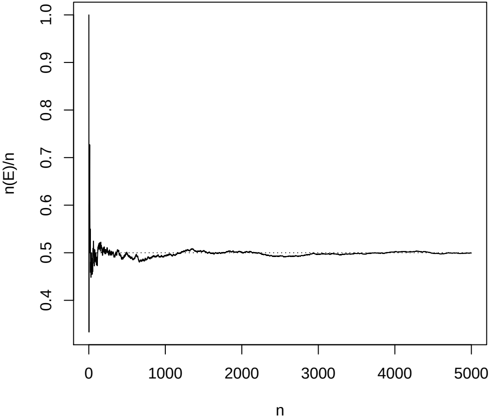
Mathematically, probability $P(E)$ is a number between 0 and 1 that is assigned to the event E from a random experiment.

How to assign probabilities?

- Subjective probability: a "degree of belief." (e.g., There is a 50% chance that I will study tonight.")
- Relative frequency probability: based on how often an event occurs over a very large sample space; i.e., $P(E) = \lim_{n\to\infty} n(A)/n$.
- Equally-likely rule: proability of each member of the sample space is the same.
- ...

Relative frequency probability

Flip a fair coin repeatedly, the relative frequency of observing "head" approaches the probability $P(''head'') = 0.5$.



However, using this to assign probability is NOT applicable in real applications. This is merely for interpretation.

Random: Equally-likely Outcomes

Whenever a sample space consists of N possible outcomes that are equally likely, the probability of each outcome is $1/N$.

Example

In a batch of 100 diodes, 1 is laser diode. A diode is **randomly** selected from the batch. **Random** means each diode has an equal chance of being selected. The probability of choosing the laser diode is $1/100$ or 0.01, because each outcome in the sample space is equally likely.

Example

Again, from a bin of 50 parts, 6 parts are selected **randomly** without replacement. The bin contains 3 defective parts and 47 nondefective parts. What is the probability that exactly 2 defective parts are selected in the sample?

Answer: when **randomly** appears, it means equally-likely rule!

$$
\begin{aligned}
& P(\text{exactly 2 defective parts}) \\
= & \frac{\#\text{ of ways to select 6 parts of which 2 are defective}}{\#\text{ of ways to select 6 parts}} \\
= & \frac{C_2^3 C_4^{47}}{C_6^{50}} \\
= & \frac{nCr(3,2) \times nCr(47,4)}{nCr(50,6)} \\
= & \, 0.03367347
\end{aligned}
$$

Probability of an Event (Discrete)

We now restrict our attention to a discrete sample space. By discrete, it means the sample sapce may be

- A finite set of outcomes; (e.g., number of winnings Gamecock can achieve in the next season)
- A countably infinite set of outcomes. (e.g., number of emails one receives on one day)

For a discrete sample space, the probability of an event E equals the sume of the probabilities of the outcomes in E.

Example

A random experiment has a sample space $S = \{a, b, c, d\}$. These outcomes are not equally-likely; their probabilities are: $0.1, 0.3, 0.5, 0.1$. Let event $A = \{a, b\}$, $B = \{b, c, d\}$, and $C = \{d\}$. Then

- $P(A) = P(a) + P(b) = 0.1 + 0.3 = 0.4$
- $P(B) = 0.3 + 0.5 + 0.1 = 0.9$
- $P(C) = P(d) = 0.1$
- $P(A^c) = P(\{c, d\}) = P(c) + P(d) = 0.5 + 0.1 = 0.6 = 1 - P(A)$; $P(B^c) = 1 - P(B) = 0.1$; $P(C^c) = 1 - 0.1 = 0.9$.
- $P(A \cap B) = P(b) = 0.3$,
  $P(A \cup B) = P(\{a, b, c, d\}) = P(S) = 1$, and
  $P(A \cap C) = P(\emptyset) = 0$.

We observe $P(S) = 1$, $P(\emptyset) = 0$, $P(A^c) = 1 - P(A)$.

Example

A wafer is randomly selected from a batch that is classified by contamination and location.

	Location in Sputtering Tool		
Contamination	Center	Edge	Total
Low	514	68	582
High	112	246	358
Total	626	314	

Let H be the event of high concentrations of contaminants. Let C be the event of the wafer being located at the center of a sputtering tool.

- $P(H) = 358/940$
- $P(C) = 626/940$
- $P(H \cap C) = 112/940$
- $P(H \cup C) = (358 + 626 - 112)/940 = P(H) + P(C) - P(H \cap C)$

Axioms of Probability

The assignment of probability to events from a random experiment must satisfy the following properties:

Axioms

If S is the sample space and E is any event from the random experiment,

1. $P(S) = 1$
2. $0 \leq P(E) \leq 1$ (0 means impossible; 1 mean certainty)
3. For any two events E_1 and E_2 with $E_1 \cap E_2 = \emptyset$ (mutually exclusive),
$$P(E_1 \cup E_2) = P(E_1) + P(E_2)$$

The axioms imply that

- $P(\emptyset) = 0$ and $P(E^c) = 1 - P(E)$
- If $E_1 \subset E_2$, then $P(E_1) \leq P(E_2)$.

Addition Rules

For any two events A and B, the probability of union is given by

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

If A and B are mutually exclusive, then $P(A \cap B) = P(\emptyset) = 0$ and

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

Addition Rules: 3 or more events

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

If a collection of events E_i are pairwise mutually exclusive; i.e., $E_i \cap E_j = \emptyset$ for $i \neq j$, then

$$P(E_1 \cup E_2 \cup \cdots \cup E_k) = \sum_{i=1}^{k} P(E_i).$$

Example

Let X denote the pH of a sample. Consider the event that X is greater than 6.5 but less than or equal to 7.8. Then $P(6.5 < X \leq 7.8) = P(6.5 < X \leq 7) + P(7 < X \leq 7.5) + P(7.5 < X \leq 7.8)$.

Conditional Probability

$P(B|A)$ is the probability of event B occurring, given that event A has already occurred.

		Surface Flaws		
		Yes (event F)	No	Total
Defective	Yes (event D)	10	18	28
	No	30	342	372
	Total	40	360	400

We have 400 parts classified by surface flaws and as (functionally) defective.
Let D denote the event that a part is defective, and
F the event that a part has a surface flaw.
The probability of D given that a part has a flaw, as $P(D|F)$.
25% of the parts with flaws are defective, $P(D|F) = 0.25$.
5% of the parts without flaws are defective, $P(D|F^c) = 0.05$.
What are $P(D^c|F)$ and $P(D^c|F^c)$?

Conditional Probability Rule and Multiplication Rule

The conditional probability of an event B given an event A, denoted as $P(B|A)$, is:

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

Consequently, we have the Multiplication Rule:

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

Example

A batch of 50 parts contains 10 made by Tool 1 and 40 made by Tool 2. If 2 parts are selected **randomly**.

(a) What is the probability that the 1st part came from Tool 1 and the 2nd part came from Tool 2?

(b) What is the probability that the 2nd part came from Tool 2, given that the 1st part came from Tool 1?

Answer: Let E_1 denote the event that the 1st part came from Tool 1; E_2 the 2nd part came from Tool 2.
(a): $P(E_2 \cap E_1) = \frac{10}{50} \times \frac{40}{49} = 8/49$
(b): $P(E_2|E_1) = P(E_2 \cap E_1)/P(E_1) = (8/49)/(10/50) = 40/49$, where $P(E_1) = 10/50$.

Example

The probability that the first stage of a numerically controlled machining operation for high-rpm pistons meets specifications is 0.90. Failures are due to metal variations, fixture alignment, cutting blade condition, vibration, and ambient environmental conditions. Given that the first stage meets specifications, the probability that a second stage of machining meets specifications is 0.95. What is the probability that both stages meet specifications?

Answer: Let A and B denote the events that the first and second stages meet specifications, respectively. The probability requested is

$$P(A \cap B) = P(B|A)P(A) = 0.95 * 0.9 = 0.855.$$

Although it is also true that $P(A \cap B) = P(A|B)P(B)$, the information provided in the problem does not match this second formulation.

Total Probability Rule

For any two events A and B:
$P(B) = P(B \cap A) + P(B \cap A^c) = P(B|A)P(A) + P(B|A^c)P(A^c)$.

For more than 2 events:
Assume $E_1, E_2, \ldots, E_k$ are k mutually exclusive and exhaustive sets; i.e.,

- $E_i \cap E_j = \emptyset$ for $i \neq j$ (mutually exclusive)
- $E_1 \cup E_2 \cup \cdots \cup E_k = S$ (exhaustive)

Then

$$\begin{aligned} P(B) &= P(B \cap E_1) + P(B \cap E_2) + \cdots + P(B \cap E_k) \\ &= P(B|E_1)P(E_1) + P(B|E_2)P(E_2) + \cdots + P(B|E_k)P(E_k). \end{aligned}$$

Example

Let F denote the event that the product fails, and H the event that the chip is exposed to high levels of contamination. Find $P(F)$.

Probability of Failure	Level of Contamination	Probability of Level
0.1	High	0.2
0.005	Not high	0.8

Answer: The third column tells us that $P(H) = 0.2$ and $P(H^c) = 0.8$. The first column tells $P(F|H) = 0.1$ and $P(F|H^c) = 0.005$. We can use total probability rule to find $P(F)$:

$$
\begin{aligned}
P(F) =& P(F|H)P(H) + P(F|H^c)P(H^c) \\
=& 0.1 \times 0.2 + 0.005 \times 0.8 = 0.024.
\end{aligned}
$$

Example

Find $P(F)$ based on the following information.

Probability of Failure	Level of Contamination	Probability of Level
0.100	High	0.2
0.010	Medium	0.3
0.001	Low	0.5

Answer: The third column tells us that $P(H) = 0.2$, $P(M) = 0.3$ and $P(L) = 0.5$. We see that H, M, L are mutually exclusive and $P(H) + P(M) + P(L) = 1$ indicating they are also exhaustive. The first column tells $P(F|H) = 0.1$, $P(F|M) = 0.01$, and $P(F|L) = 0.001$. We can use total probability rule to find $P(F)$:

$$
\begin{aligned}
P(F) =& P(F|H)P(H) + P(F|M)P(M) + P(F|L)P(L) \\
=& 0.1 \times 0.2 + 0.01 \times 0.3 + 0.001 \times 0.5 \\
=& 0.0235.
\end{aligned}
$$

Table 1 provides an example of 400 parts classified by surface flaws and as (functionally) defective. Suppose that the situation is different and follows Table 2. Let F denote the event that the part has surface flaws. Let D denote the event that the part is defective.

TABLE 1 Parts Classified				TABLE 2 Parts Classified (data chg'd)			
	Surface Flaws				Surface Flaws		
Defective	Yes (F)	No (F')	Total	Defective	Yes (F)	No (F')	Total
Yes (D)	10	18	28	Yes (D)	2	18	20
No (D')	30	342	372	No (D')	38	342	380
Total	40	360	400	Total	40	360	400
	$P(D\|F)=$	10/40 =	0.25		$P(D\|F)=$	2/40 =	0.05
	$P(D)=$	28/400 =	0.10		$P(D)=$	20/400 =	0.05
			not same				same
	Events D & F are **dependent**				Events D & F are **independent**		

Independence

Two events are independent if any one of the following equivalent statements is true:

1. $P(A|B) = P(A)$
2. $P(B|A) = P(B)$
3. $P(A \cap B) = P(A) \cdot P(B)$

This means that occurrence of one event has on impact on the probability of occurrence of the other event.

- If A and B are mutually exclusive, are they independent?
- If (A and B) are independent, so are (A and B^c), (A^c and B), (A^c and B^c).

Independence with multiple events

The events $E_1, E_2, \dots, E_n$ are independent, if and only if, for any subsets of these events:

$$P(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_k}) = P(E_{i1}) \cdot P(E_{i_2}) \cdots P(E_{i_k}).$$

Circuit Operation

The following circuit operates only if there is a path of functional devices from left to right. The probability that each device functions is shown on the graph. Assume that devices fail **independently**. What is the probability that the circuit operates?

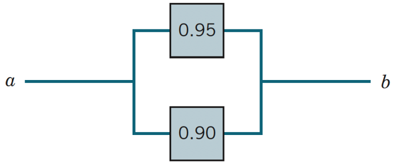
Answer: The circuit operates if an only if the two parts operate together.

$$P(L \cap R) = P(L) \cdot P(R) = 0.8 \times 0.9 = 0.72.$$

Practical Interpretation: Notice that the probability that the circuit operates degrades to approximately 0.7 when all devices are required to be functional. The probability that each device is functional needs to be large for a circuit to operate when many devices are connected in series.

Circuit Operation

Assume that devices fail **independently**. What is the probability that the circuit operates?



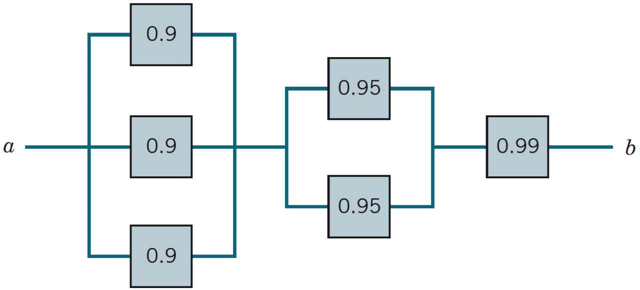
Answer: The circuit operates if at least one device operates.

$$
\begin{aligned}
P(T \cup B) =& 1 - P\{(T \cup B)^c\} = 1 - P(T^c \cap B^c) \\
=& 1 - P(T^c)P(B^c) = 1 - (1 - 0.95)(1 - 0.9) = 0.995
\end{aligned}
$$

Practical Interpretation: Notice that the probability that the circuit operates is larger than the probability that either device is functional. This is an advantage of a parallel architecture.

Circuit Operation

Assume that devices fail **independently**. What is the probability that the circuit operates?



Answer:

$$\begin{aligned} P(L \cap M \cap R) &= P(L)P(M)P(R) \\ &= (1 - 0.1^3)(1 - 0.05^2)0.99 = 0.987. \end{aligned}$$

Bayes' Theorem

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \text{ for } P(B) > 0.$$

Example

Let F denote the event that the product fails, and H the event that the chip is exposed to high levels of contamination. Find $P(H|F)$, the conditional probability that a high level of contamination was present when a failure occurred is to be determined.

Probability of Failure	Level of Contamination	Probability of Level
0.1	High	0.2
0.005	Not high	0.8

$$P(H|F) = \frac{P(F|H)P(H)}{P(F)} = \frac{0.1 \cdot 0.2}{0.24} = 0.83.$$

Example: Medical Diagnostic

Because a new medical procedure has been shown to be effective in the early detection of an illness, a medical screening of the population is proposed. The probability that the test correctly identifies someone with the illness as positive (known as the sensitivity) is 0.95, and the probability that the test correctly identifies someone without the illness as negative (known as the specificity) is 0.99. The incidence of the illness in the general population is 0.0001. You take the test, and the result is positive. What is the probability that you have the illness?

Answer: Let I denote the event that you have the illness, and let T denote the event that the test signals positive. Then $P(T|I) = 0.95$, $P(T^c|I^c) = 0.99$, and $P(I) = 0.0001$.

$$
\begin{aligned}
P(I|T) &= \frac{P(T|I)P(I)}{P(T|I)P(I) + P(T|I^c)P(I^c)} \\
&= \frac{0.95(0.0001)}{0.95(0.0001) + (1-0.99)(1-0.0001)} = 0.0094
\end{aligned}
$$

Bayes' Theorem with total probability rule

If $E_1, E_2, \ldots, E_k$ are k mutually exclusive and exhausive events and B is any event with $P(B) > 0$, then

$$
\begin{aligned}
P(E_1|B) &= \frac{P(B|E_1)P(E_1)}{P(B)} \\
&= \frac{P(B|E_1)P(E_1)}{P(B|E_1)P(E_1) + P(B|E_2)P(E_2) + \cdots + P(B|E_k)P(E_k)}.
\end{aligned}
$$

Example: Bayesian Network

A printer manufacturer obtained the following three types of printer failure probabilities. Hardware $P(H) = 0.3$, software $P(S) = 0.6$, and other $P(O) = 0.1$. Also, $P(F|H) = 0.9$, $P(F|S) = 0.2$, and $P(F|O) = 0.5$. If a failure occurs, determine if it's most likely due to hardware, software, or other.

Answer: We need to find out which of $P(H|F)$, $P(S|F)$, $P(O|F)$ is the largest. We also note H, S, O are mutually exclusive and exhaustive events.

$$P(H|F) = \frac{P(F|H)P(H)}{P(F)} = \frac{0.9 \cdot 0.3}{0.44} = 0.6136.$$

where $P(F) = P(F|H)P(H) + P(F|S)P(S) + P(F|O)P(O) = 0.9(0.3) + 0.2(0.6) + 0.5(0.1) = 0.44$. Similarly, $P(S|F) = 0.12/0.44 = 0.2727$ and $P(O|F) = 0.05/0.44 = 0.1136$.