P1 - Probability

STAT 587 (Engineering) - Iowa State University

January 23, 2019

Interpretation

What do we mean when we say the word probability? For example,

- The probability of a recession in the next 12 months is 0.40%.
- The probability the Chiefs will beat the Patriots is 68%.
- The probability the government shutdown will end in the next two weeks is 30%.

Interpretations:

- Relative frequency: Probability is the proportion of times the event occurs as the number of times the event is attempt tends to infinity.
- Personal belief: Probability is a statement about your personal belief in the event occuring.

Probability

Example

- Consider the event C: a successful connection to the internet from a laptop.
- From our experience with the wireless network and our internet service provider, we believe the probability we successfully connect is 90 %.
- We write: $P(C) = 0.9$
- To be able to work with probabilities, in particular, to be able to compute probabilities of events, a mathematical foundation is necessary.

Set theory

Definition

A set is a collection of things. We use the following notation

- $\omega \in A$ means ω is an element of the set A,
- $\omega \notin A$ means ω is not an element of the set A,
- $A \subseteq B$ (or $B \supseteq A$) means the set A is a subset of B (with the sets possibly being equal), and
- $A \subset B$ (or $B \supset A$) means the set A is a proper subset of B, i.e. there is at least one element in B that is not in A.

The sample space, Ω, is the set of all outcomes of an experiment.

Examples

Example

The set of all possible sums of two 6-sided dice rolls is
$\Omega = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ and

- $2 \in \Omega$
- $1 \notin \Omega$
- $\{2, 3, 4\} \subset \Omega$

Set comparison, operations, terminology

For the following $A, B \subseteq \Omega$ where Ω is the implied universe of all elements under study,

1. Union ($\cup$): A union of events is an event consisting of all the outcomes in these events.

$$A \cup B = \{\omega \mid \omega \in A \text{ or } \omega \in B\}$$

2. Intersection ($\cap$): An intersection of events is an event consisting of the common outcomes in these events.

$$A \cap B = \{\omega \mid \omega \in A \text{ and } \omega \in B\}$$

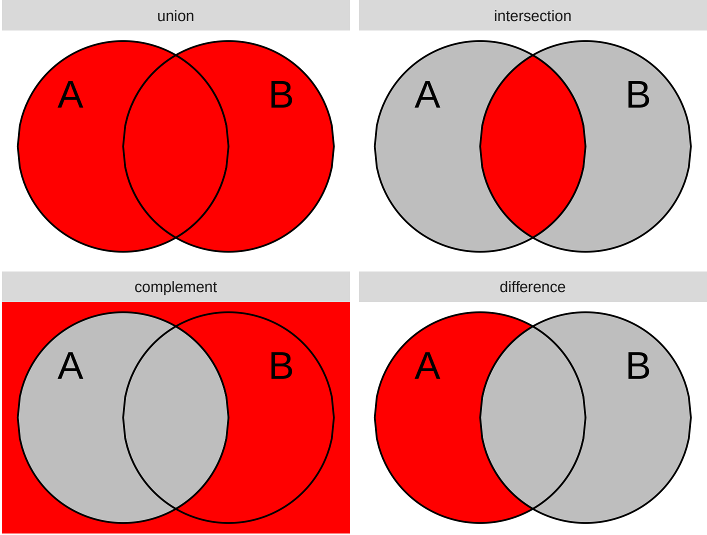
3. Complement (A^C): A complement of an event A is an event that occurs when event A does not happen.

$$A^C = \{\omega \mid \omega \notin A \text{ and } \omega \in \Omega\}$$

4. Set difference ($A \setminus B$): All elements in A that are not in B, i.e.

$$A \setminus B = \{\omega | \omega \in A \text{ and } \omega \notin B\}$$

Venn diagrams



Example

Consider the set Ω equal to all possible sum of two 6-sided die rolls i.e. $\Omega = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ and two subsets

- all odd rolls: $A = \{3, 5, 7, 9, 11\}$
- all rolls below 6: $B = \{2, 3, 4, 5\}$

Then we have

- $A \cup B = \{2, 3, 4, 5, 7, 9, 11\}$
- $A \cap B = \{3, 5\}$
- $A^C = \{2, 4, 6, 8, 10, 12\}$
- $B^C = \{6, 7, 8, 9, 10, 11, 12\}$
- $A \setminus B = \{7, 9, 11\}$
- $B \setminus A = \{2, 4\}$

Set comparison, operations, terminology (cont.)

5. Empty Set $\emptyset$ is a set having no elements, i.e. $\{\}$. The empty set is a subset of every set:
$$\emptyset \subseteq A$$
6. Disjoint sets: Sets A, B are disjoint if their intersection is empty:
$$A \cap B = \emptyset$$
7. Mutually exclusive sets: Sets $A_1, A_2, \ldots$ are mutually exclusive if all pairs of these events are disjoint:
$$A_i \cap A_j = \emptyset \text{ for any } i \neq j$$
8. De Morgan's Laws:
$$(A \cup B)^C = A^C \cap B^C \qquad \text{and} \qquad (A \cap B)^C = A^C \cup B^C$$

Examples

Let $A = \{2, 3, 4\}, B = \{5, 6, 7\}, C = \{8, 9, 10\}, D = \{11, 12\}$. Then

- $A \cap B = \emptyset$
- A, B, C, D are mutually exclusive
- De Morgan's:

$$\begin{aligned} (A \cup B) &= \{2, 3, 4, 5, 6, 7\} \\ (A \cup B)^C &= \{8, 9, 10, 11, 12\} \end{aligned}$$

$$\begin{aligned} A^C &= \{5, 6, 7, 8, 9, 10, 11, 12\} \\ B^C &= \{2, 3, 4, 8, 9, 10, 11, 12\} \\ A^C \cap B^C &= \{8, 9, 10, 11, 12\} \end{aligned}$$

so, by example,

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

Kolmogorov's Axioms

To be able to work with probabilities properly - to compute with them - one must lay down a set of postulates.

A system of probabilities (a probability model) is an assignment of numbers $P(A)$ to events $A \subseteq \Omega$ such that

(i) $0 \leq P(A) \leq 1$ for all A

(ii) $P(\Omega) = 1$.

(iii) if $A_1, A_2, \ldots$ are (possibly, infinite many) mutually exclusive events (i.e. $A_i \cap A_j = \emptyset$ for all $i \neq j$) then

$$P(A_1 \cup A_2 \cup \ldots) = P(A_1) + P(A_2) + \ldots = \sum_i P(A_i).$$

Kolmogorov's Axioms (cont.)

These are the basic rules of operation of a probability model

- every valid model must obey these,
- any system that does, is a valid model.

Whether or not a particular model is realistic is different question.

Example: Draw a single card from a standard deck of playing cards:
$\Omega = \{red, black\}$ Two different, equally valid probability models are:

Model 1	Model 2
$P(\Omega) = 1$	$P(\Omega) = 1$
$P(red) = 0.5$	$P(red) = 0.3$
$P(black) = 0.5$	$P(black) = 0.7$

Mathematically, both schemes are equally valid. But, of course, our real world experience would prefer model 1 over model 2 as the 'correct' model.

Useful Consequences of Kolmogorov's Axioms

Let A, $B \subseteq \Omega$.

- Probability of the Complementary Event: $P\left(A^C\right) = 1 - P(A)$
  Corollary: $P(\emptyset) = 0$
- Addition Rule of Probability

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

- If $A \subseteq B$, then $P(A) \leq P(B)$.

Example: Using Kolmogorov's Axioms

We attempt to access the internet from a laptop at home. We connect successfully if and only if the wireless (WiFi) network works *and* the internet service provider (ISP) network works. Assume

$$\begin{aligned} P(\text{ WiFi up }) &= .9 \\ P(\text{ ISP up }) &= .6, \text{ and} \\ P(\text{ WiFi up and ISP up }) &= .55. \end{aligned}$$

1. What is the probability that the WiFi is up or the ISP is up?
2. What is the probability that both the WiFi and the ISP are down?
3. What is the probability that we fail to connect?

Solution

Let $A \equiv$ WiFi up; $B \equiv$ ISP up

1. What is the probability that the WiFi is up or the ISP is up?

$$P(\text{ WiFi up or ISP up}) = P(A \cup B) = 0.9 + 0.6 - 0.55 = 0.95$$

2. What is the probability that both the WiFi and the ISP are down?

$$\begin{aligned} P(\text{ WiFi down and ISP down}) &= P\left(A^C \cap B^C\right) = P\left([A \cup B]^C\right) \\ &= 1 - .95 = .05 \end{aligned}$$

3. What is the probability that we fail to connect?

$$\begin{aligned} &P(\text{ WiFi down or ISP down}) \\ &\quad = P\left(A^C \cup B^C\right) = P\left(A^C\right) + P\left(B^C\right) - P\left(A^C \cap B^C\right) \\ &\quad = P\left(A^C \cup B^C\right) = (1 - .9) + (1 - .6) - .05 = .1 + .4 - .05 = .45 \end{aligned}$$

Conditional probability

Definition

The conditional probability of an event A given an event B is

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

if $P(B) > 0$.

Intuitively, the fraction of outcomes in B that are also in A.

Corollary

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

Random CPUs

A box has 500 CPUs with a speed of 1.8 GHz and 500 with a speed of 2.0 GHz. The numbers of good (G) and defective (D) CPUs at the two different speeds are as shown below.

	1.8 GHz	2.0 GHz	Total
G	480	490	970
D	20	10	30
Total	500	500	1000

We select a CPU at random and observe its speed. What is the probability that the CPU is defective given that its speed is 1.8 GHz?
Let

- D be the event the CPU is defective and
- S be the event the CPU speed is 1.8 GHz.

Then

- $P(S) = 500/1000 = 0.5$
- $P(S \cap D) = 20/1000 = 0.02.$
- $P(D|S) = P(S \cap D)/P(S) = 0.02/0.5 = 0.04.$

Statistical independence

Definition

Events A and B are statistically independent if

$$P(A \cap B) = P(A) \times P(B)$$

or, equivalently,

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

Intuitively, the occurrence of one event does not affect the probability of the other.

Example

If I toss a fair coin and it comes up tails, does that affect the probability the next coin flip is heads?

WiFi example

In trying to connect my laptop to the internet, I need

- my WiFi network to be up (event A) and
- the ISP network to be up (event B).

As previous, assume the probability the WiFi network is up is 0.6 and the ISP network is up is 0.9. If the two events are independent, what is the probability we can connect to the internet?

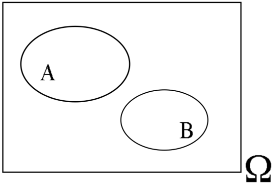
Since we have independence, we know

$$P(A \cap B) = P(A) \times P(B) = 0.6 \times 0.9 = 0.54.$$

Independence and disjoint

Warning: Independence and disjointedness are two very different concepts!

Disjointedness:

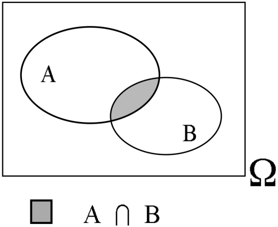


A, B are disjoint

If A and B are disjoint, their intersection is empty and therefore has probability 0:

$$P(A \cap B) = P(\emptyset) = 0.$$

Independence:



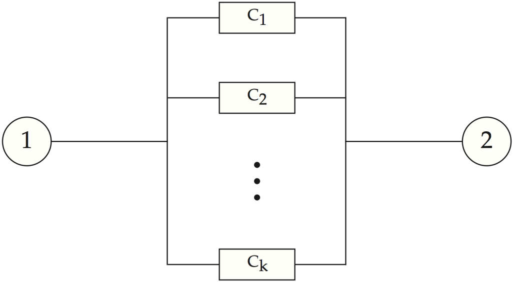
If A and B are independent events, the probability of their intersection can be computed as the product of their individual probabilities:

$$P(A \cap B) = P(A) \cdot P(B)$$

Parallel system

Definition

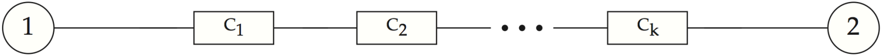
A parallel system consists of K components $c_1, \ldots, c_K$ arranged in such a way that the system works if **at least one** of the K components functions properly.



Serial system

Definition

A serial system consists of K components $c_1, \ldots, c_K$ arranged in such a way that the system works if and only if **all** of the components function properly.



Reliability

Definition

Reliability of a system is the probability the system works.

Example

The reliability of the WiFi-ISP network (assuming independence) is 0.54

Reliability of parallel systems with independent components

Let $c_1, \ldots, c_K$ denote the K components in a parallel system. Assume the K components operate independently and $P(c_k \text{ works}) = p_k$. What is the reliability of the system?

$$
\begin{aligned}
P(\text{ system works }) &= P(\text{ at least one component works }) \\
&= 1 - P(\text{ all components fail }) \\
&= 1 - P(c_1 \text{ fails and } c_2 \text{ fails } \ldots \text{ and } c_k \text{ fails }) \\
&= 1 - \prod_{k=1}^{K} P(c_k \text{ fails}) \\
&= 1 - \prod_{k=1}^{K} (1 - p_k).
\end{aligned}
$$

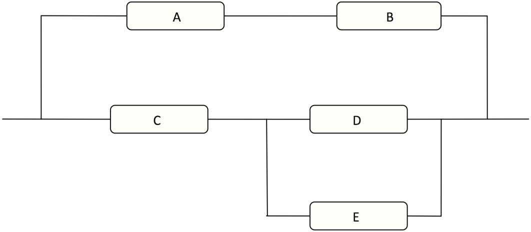
Reliability of serial systems with independent components

Let $c_1, \ldots, c_K$ denote the K components in a serial system. Assume the K components operate independently and $P(c_k \text{ works }) = p_k$. What is the reliability of the system?

$$\begin{aligned} P(\text{ system works }) &= P(\text{ all components work}) \\ &= \prod_{k=1}^{K} P(c_k \text{ works}) \\ &= \prod_{k=1}^{K} p_k. \end{aligned}$$

Reliability example

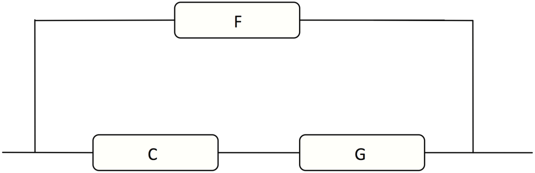
Each component in the system shown below is opearable with probability 0.92 independently of other components. Calculate the reliability.



1. Serial components A B can be replaced by a component F that operates with probability $P(A \cap B) = (0.92)^2 = 0.8464$.
2. Parallel components D and E can be replaced by component G that operates with probability $P(D \cup E) = 1 - (1 - 0.92)^2 = 0.9936$.

Reliability example (cont.)

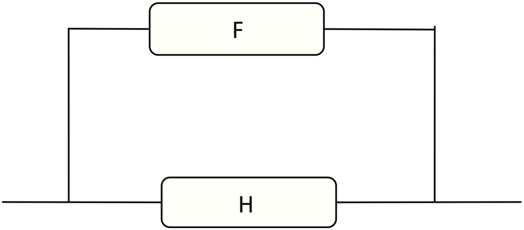
Updated circuit:



3. Serial components C and G connected can be replaced by a component H that operates with probability $P(C \cap G) = (0.92)(0.9936) = 0.9141$.

Reliability example (cont.)

Updated circuit:



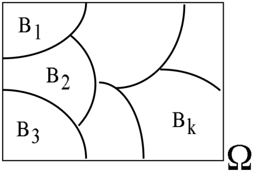
4. Parallel componenents F and H are in parallel, so the reliability of the system is $P(F \cup H) = 1 - (1 - 0.8424)(1 - 0.9141) = 0.9868$.

Partition

Definition

A collection of events $B_1, \dots B_K$ is called a partition (or cover) of Ω if

- the events are mutually exclusive (i.e., $B_i \cap B_j = \emptyset$ for $i \neq j$), and
- the union of the events is Ω (i.e., $\bigcup_{k=1}^{K} B_k = \Omega$).



Example

Consider the sum of two 6-sided die, i.e.

$$\Omega = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}.$$

Here are some covers:

- $\{2, 3, 4\}, \{5, 6, 7, 8, 9, 10, 11, 12\}$
- $\{2, 3, 4\}, \{5, 6, 7\}, \{8, 9, 10\}, \{11, 12\}$
- $A_2, A_3, \ldots, A_{12}$ where $A_i = \{i\}$
- any A and A^C where $A \subseteq \Omega$

Law of Total Probability

Theorem (Law of Total Probability)

If the collection of events $B_1, \ldots, B_K$ is a partition of Ω, and A is an event, then

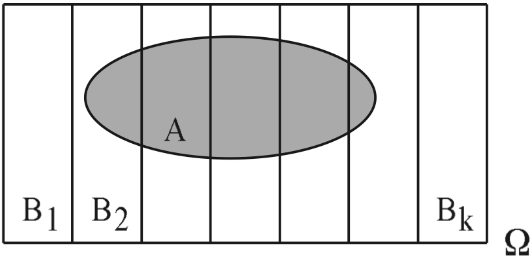
$$P(A) = \sum_{k=1}^{K} P(A|B_k)P(B_k).$$

Proof.

$$\begin{aligned}
P(A) &= P\left(\bigcup_{k=1}^{K} A \cap B_k\right) && B_1, \ldots, B_K \text{ is a partition} \\
&= \sum_{k=1}^{K} P(A \cap B_k) && A \cap B_1, \ldots, A \cap B_K \text{ are disjoint} \\
&= \sum_{k=1}^{K} P(A|B_k)P(B_k) && \text{definition of conditional probability}
\end{aligned}$$

□

Law of Total Probability graphic



Example

In the come out roll of craps, you win if the roll is a 7 or 11. By the total law of probability, the probability you win is

$$P(\text{Win}) = \sum_{i=2}^{12} P(\text{Win}|i)P(i) = P(7) + P(11)$$

since $P(\text{Win}|i) = 1$ if $i = 7, 11$ and 0 otherwise.

Bayes' rule

Theorem (Bayes' Rule)

If $B_1, \ldots, B_K$ is a partition of Ω, and A is an event, then

$$P(B_k|A) = \frac{P(A|B_k)P(B_k)}{\sum_{k=1}^K P(A|B_k)P(B_k)}.$$

Proof.

$$\begin{aligned} P(B_k|A) &= \frac{P(A \cap B_k)}{P(A)} && \text{by definition of conditional probability} \\ &= \frac{P(A|B_k)P(B_k)}{P(A)} && \text{by definition of conditional probability} \\ &= \frac{P(A|B_k)P(B_k)}{\sum_{k=1}^K P(A|B_k)P(B_k)} && \text{by Law of Total Probability} \end{aligned}$$

$\square$

Example

If you win on a come-out roll in craps, what is the probability you rolled a 7?

$$\begin{aligned} P(7|\text{Win}) &= \frac{P(\text{Win}|7)P(7)}{\sum_{i=2}^{12} P(\text{Win}|i)P(i)} \\ &= \frac{P(7)}{P(7)+P(11)}. \end{aligned}$$

CPU testing

Example

A given lot of CPUs contains 2% defective CPUs. Each CPU is tested before delivery. However, the tester is not wholly reliable:

$$\begin{aligned} P(\text{ tester says CPU is good} \mid \text{CPU is good }) &= 0.95 \\ P(\text{ tester says CPU is defective} \mid \text{CPU is defective }) &= 0.94 \end{aligned}$$

If the test device says the CPU is defective, what is the probability that the CPU is actually defective?

CPU testing (cont.)

Let

- C_g (C_d) be the event the CPU is good (defective)
- T_g (T_d) be the event the tester says the CPU is good (defective)

We know

- $0.02 = P(C_d) \quad = 1 - P(C_g)$
- $0.95 = P(T_g|C_g) = 1 - P(T_d|C_g)$
- $0.94 = P(T_d|C_d) = 1 - P(T_g|C_d)$

Using Bayes' Rule, we have

$$\begin{aligned} P(C_d|T_d) &= \frac{P(T_d|C_d)P(C_d)}{P(T_d|C_d)P(C_d)+P(T_d|C_g)P(C_g)} \\ &= \frac{P(T_d|C_d)P(C_d)}{P(T_d|C_d)P(C_d)+[1-P(T_g|C_g)][1-P(C_d)]} \\ &= \frac{0.94\times 0.02}{0.94\times 0.02+[1-0.95]\times[1-0.02]} \\ &= 0.28 \end{aligned}$$