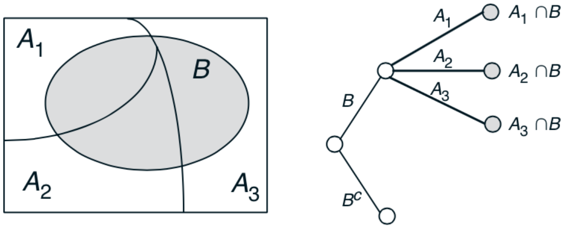
Total probability - Bayes' formula

January 14, 2021

Partition

$A_1, \ldots, A_n$ is a partition of Ω if

- mutually exclusive: $A_i A_j = \emptyset$ for $i \neq j$
- $A_1 \cup A_2 \cup \ldots A_n = \Omega$

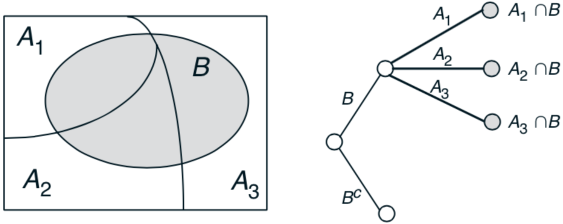


$BA_1, \ldots, BA_n$ is a partition of B

Partition

$A_1, \ldots, A_n$ is a partition of Ω if

- mutually exclusive: $A_i A_j = \emptyset$ for $i \neq j$
- $A_1 \cup A_2 \cup \ldots A_n = \Omega$



$BA_1, \ldots, BA_n$ is a partition of B

Total probability formula - divide - and - conquer

- Partition sample space into $A_1, A_2, \ldots, A_n$
- Know $P(B|A_i)$ for every i
- Compute $P(B)$

$$P(B) = \sum_{i=1}^{n} P(BA_i) = \sum_{i=1}^{n} P(B|A_i)P(A_i).$$

Example

In semiconductor manufacturing the probability is 0.10 that a chip that is subjected to high levels of contamination during manufacturing causes a product failure. The probability is 0.005 that a chip that is not subjected to high contamination levels during manufacturing causes a product failure. In a particular production run, 20% of the chips are subject to high levels of contamination. What is the probability that a product using one of these chips fails?

Solution

Summarize information

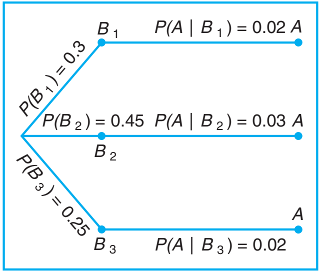
Level of contamination	Prob of level	Prob of **failure** (F) subject to level
Hight (H)	$P(H) = 0.2$	$P(F\|H) = 0.1$
Not hight (H^c)	$P(H^c) = 0.8$	$P(F\|H^c) = 0.005$

$$P(F) = P(FH) + P(FH^c)$$
$$= P(H)P(F|H) + P(H^c)P(F|H^c) = 0.024$$

Example

In a certain assembly plant, three machines, B_1, B_2, and B_3, make 30%, 45%, and 25%, respectively, of the products. It is known from past experience that 2%, 3%, and 2% of the products made by each machine, respectively, are defective. Now, suppose that a finished product is randomly selected. What is the probability that it is defective?

Solution



- $P(AB_1) = P(B_1)P(A|B_1) = (.3)(.02) = .006$
- $P(AB_2) = P(B_2)P(A|B_2) = (.45)(.03) = .0135$
- $P(AB_3) = P(B_1)P(A|B_1) = (.25)(.02) = .005$
- $P(A) = P(AB_1) + P(AB_2) + P(AB_3) = 0.0245$

Re-evaluate

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

Solution

$$P(B_3|A) = \frac{P(B_3A)}{P(A)}$$
$$= \frac{0.005}{0.0245} = \frac{10}{49}$$

Bayes' Formula

- Prior probability $P(A_i)$ - initial belief
- Know $P(B|A_i)$ for each i
- Given B occurs, wish to revise (update) "belief" $P(A_i|B)$

$$\boxed{P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{k=1}^{n} P(B|A_i)P(A_i)}}$$

Bayes' rule is often used for **inference**. There are a number of “causes” that may result in a certain “effect.” We observe the effect, and we wish to infer the cause.

Example

A manufacturing firm employs three analytical plans for the design and development of a particular product. For cost reasons, all three are used at varying times. In fact, plans 1, 2, and 3 are used for 30%, 20%, and 50% of the products, respectively.

The defect rate is different for the three procedures as follows:

$$P(D|P_1) = 0.01, P(D|P_2) = 0.03, P(D|P_3) = 0.02$$

where $P(D|P_j)$ is the probability of a defective product, given plan j.

If a random product was observed and found to be defective, which plan was most likely used and thus responsible?

Solution

- Need to compare $P(P_1|D)$, $P(P_2|D)$, $P(P_3|D)$
- Info:

	$P(P_i)$	$P(D\|P_i)$	$P(DP_i)$
P_1	0.3	0.01	0.003
P_2	0.2	0.03	0.006
P_3	0.5	0.03	0.01
			$P(D) = 0.19$

- $P(P_1|D) = \frac{P(P_1 D}{P(D)} = \frac{.003}{.19} = .158$
- Similarly, $P(P_2|D) = .316, P(P_3|D) = .526$

The conditional probability of a defect given plan 3 is the largest of the three; thus a defective for a random product is most likely the result of the use of plan 3

Practice

Printer failures are associated with three types of problems: hardware, software, and other (such as connectors), with probabilities 0.1, 0.6, and 0.3, respectively. The probability of a printer failure given a hardware problem is 0.9, given a software problem is 0.2, and given any other type of problem is 0.5. If a customer enters the manufacturer's Web site to diagnose a printer failure, what is the most likely cause of the problem?

Example

A plane is missing and it was equally likely to have gone down in any of three possible regions. Let $1 - \alpha_i$ denote the probability the plane will be found upon a search of the i-th region when the plane is, in fact, in that region, $i = 1, 2, 3$. What is the conditional probability that the plane is in the i-th region, given that a search of region 1 is unsuccessful, $i = 1, 2, 3$?

Solution

- A_i={the plane is in region i}
- B={search of region 1 was unsuccessful}
- Need $P(A_i|B) = ?$

Solution

- A_i={the plane is in region i}
- B={search of region 1 was unsuccessful}
- Need $P(A_i|B) = ?$

Solution

- A_i={the plane is in region i}
- B={search of region 1 was unsuccessful}
- Need $P(A_i|B) = ?$

Solution

Need to find

$$P(A_i B) \text{ and } P(B)$$

with information

- $P(A_i) = \frac{1}{3}$
- $P(\text{ plane is found in region } i | A_i) = 1 - \alpha$

Solution

$$P(A_1B) = ?$$

A_1B means that

- Plane is in region 1
- Search in region 1 was unsuccessful = plane was not found in region 1

$$P(A_1B) = P(A_1)P(B|A_1) = \frac{1}{3} * \alpha_1 = \frac{\alpha_1}{3}$$

Solution

$$P(A_1B) = ?$$

A_1B means that

- Plane is in region 1
- Search in region 1 was unsuccessful = plane was not found in region 1

$$P(A_1B) = P(A_1)P(B|A_1) = \frac{1}{3} * \alpha_1 = \frac{\alpha_1}{3}$$

Solution

$$P(A_1B) = ?$$

A_1B means that

- Plane is in region 1
- Search in region 1 was unsuccessful = plane was not found in region 1

$$P(A_1B) = P(A_1)P(B|A_1) = \frac{1}{3} * \alpha_1 = \frac{\alpha_1}{3}$$

Solution

A_2B means that

- Plane is in region 2
- Search in region 1 was unsuccessful = plane was not found in region 1

$$P(A_2B) = P(A_2)P(B|A_2) = \frac{1}{3} * 1 = \frac{1}{3}$$

Solution

A_2B means that

- Plane is in region 2
- Search in region 1 was unsuccessful = plane was not found in region 1

$$P(A_2B) = P(A_2)P(B|A_2) = \frac{1}{3} * 1 = \frac{1}{3}$$

Solution

A_3B means that

- Plane is in region 3
- Search in region 1 was unsuccessful = plane was not found in region 1

$$P(A_3B) = P(A_3)P(B|A_3) = \frac{1}{3} * 1 = \frac{1}{3}$$

Solution

A_3B means that

- Plane is in region 3
- Search in region 1 was unsuccessful = plane was not found in region 1

$$P(A_3B) = P(A_3)P(B|A_3) = \frac{1}{3} * 1 = \frac{1}{3}$$

Solution

$$
\begin{aligned}
P(B) &= P(A_1B) + P(A_2B) + P(A_3B) \\
&= \alpha_1 \times \frac{1}{3} + 1 \times \frac{1}{3} + 1 \times \frac{1}{3} \\
&= \frac{\alpha_1 + 2}{3}
\end{aligned}
$$

Solution

$$
\begin{aligned}
P(B) =& P(A_1B) + P(A_2B) + P(A_3B) \\
=& \alpha_1 \times \frac{1}{3} + 1 \times \frac{1}{3} + 1 \times \frac{1}{3} \\
=& \frac{\alpha_1 + 2}{3}
\end{aligned}
$$

Solution

$$P(A_1|B) = \frac{P(A_1B)}{P(B)} = \frac{\alpha_1}{\alpha_1 + 2}$$

$$P(A_2|B) = \frac{P(A_2B)}{P(B)} = \frac{1}{\alpha_1 + 2}$$

$$P(A_3|B) = \frac{P(A_3B)}{P(B)} = \frac{1}{\alpha_1 + 2}$$

Solution

$$P(A_1|B) = \frac{P(A_1B)}{P(B)} = \frac{\alpha_1}{\alpha_1 + 2}$$

$$P(A_2|B) = \frac{P(A_2B)}{P(B)} = \frac{1}{\alpha_1 + 2}$$

$$P(A_3|B) = \frac{P(A_3B)}{P(B)} = \frac{1}{\alpha_1 + 2}$$

Solution

$$P(A_1|B) = \frac{P(A_1B)}{P(B)} = \frac{\alpha_1}{\alpha_1 + 2}$$

$$P(A_2|B) = \frac{P(A_2B)}{P(B)} = \frac{1}{\alpha_1 + 2}$$

$$P(A_3|B) = \frac{P(A_3B)}{P(B)} = \frac{1}{\alpha_1 + 2}$$

Summary

- Partition $A_1, \ldots, A_n$
  - $A_i A_j = \emptyset$ for all $i \neq j$
  - $\cup_i A_i = \Omega$
- Total rule

$$P(B) = P(BA_1) + \ldots P(BA_n)$$
$$= P(A_1)P(B|A_1) + \cdots + P(A_n)P(B|A_n)$$

- Bayes's rule (update information)

$$P(A_k|B) = \frac{P(BA_k)}{P(B)}$$

$$= \frac{P(A_k)P(B|A_k)}{P(A_1)P(B|A_1) + \cdots + P(A_n)P(B|A_n)}$$