Law of Total Probability and Bayes' Rule

Let us reconsider the experiment involving colorblindness from previous Section.

Example: Suppose that in the general population, there are 51% men and 49% women, and that the proportion of colorblind men and women are shown in the probability table below:

	Men (B)	Women(B^C)	Total
Colorblind(A)	**0.04**	**0.002**	**0.042**
Not Colorblind(A^C)	**0.47**	**0.488**	**0.958**
Total	**0.51**	**0.49**	

Notice that the two events

B: the person selected is a man

B^C: the person selected is a woman

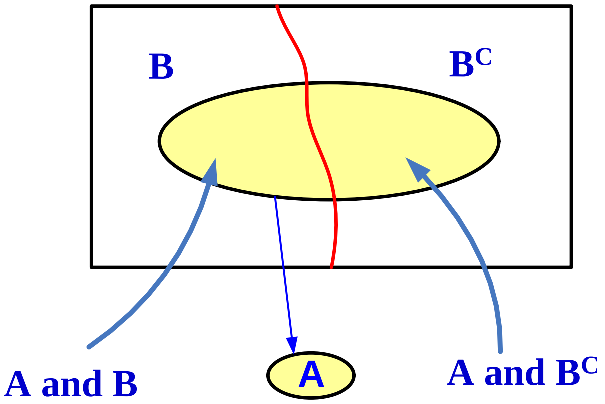
taken together make up the sample space S, consisting of both men and women .

- **Since colorblind people can be either male or female, the event A, which is that a person is colorblind, consist of those simple events that are in A and B and those simple events that are in A and B^C.**
- **Since these two intersections are mutually are exclusive , we can write the event A as**

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

and

$$P(A) = P(A \cap B) + P(A \cap B^C)$$
$$= 0.04 + 0.002 = 0.042$$



Suppose now that the sample space can be partitioned into k subpopulations, $S_1, S_2, S_3, \ldots, S_k$, that, as in the colorblindness example, are <u>mutually exclusive and exhaustive</u>*; that is taken together they make up the entire sample space .

* Collectively exhaustive events

From Wikipedia, the free encyclopedia

<u>In probability theory, a set of events is **collectively exhaustive** if at least one of the events must occur</u>. For example, when rolling a six-sided die, the outcomes 1, 2, 3, 4, 5, and 6 are collectively exhaustive, because they encompass the entire range of possible outcomes.

Another way to describe <u>collectively exhaustive events is that their union must cover all the events within the entire sample space</u>. For example, events A and B are said to be collectively exhaustive if $A \cup B = S$ where S is the sample space.

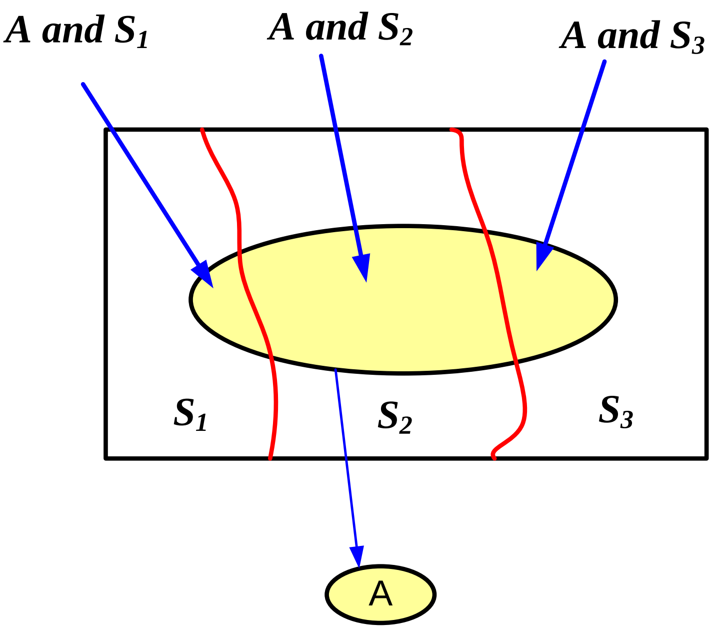
In a similar way, we can express an event A as

$$A = (A \cap S_1) \cup (A \cap S_2) \cup ... \cup (A \cap S_k)$$

Then

$$P(A)=P(A\cap S_1)+P(A\cap S_2)+P(A\cap S_3)+\ldots+P(A\cap S_k)$$

Given a set of events S_1, S_2, S_3 that are mutually exclusive and exhaustive and event A.



We can express the event A as

$$A = (A \cap S_1) \cup (A \cap S_2) \cup (A \cap S_3).$$

The probability of the event A can be expressed as

$$P(A) = P(A \cap S_1)+P(A \cap S_2)+P(A \cap S_3).$$

Law of Total Probability

Given a set of events S_1, S_2,..., S_k that are mutually exclusive and exhaustive and an event A can be expressed as

$$P(A)=P(A\cap S_1)+P(A\cap S_2)+P(A\cap S_3)+\ldots+P(A\cap S_k)$$

Or

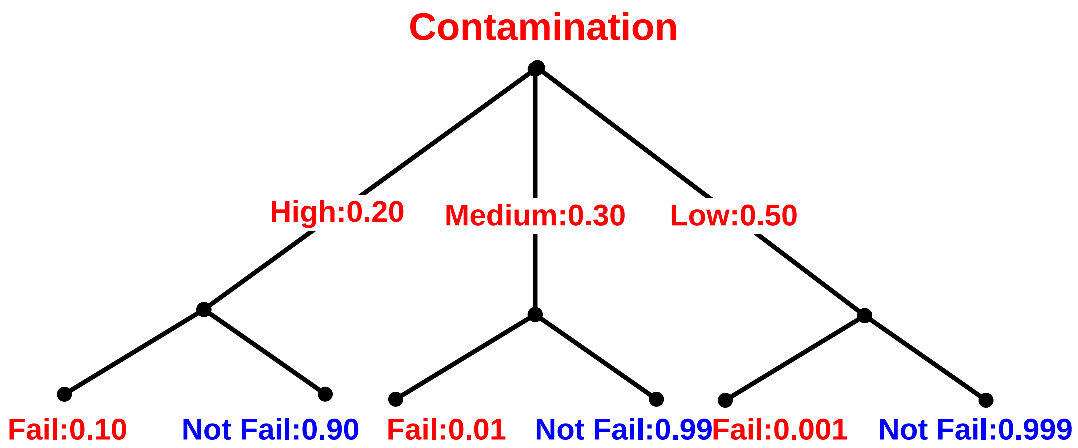
$$P(A)=P(S_1)P(A\backslash S_1)+P(S_2)P(A\backslash S_2)+\ldots+P(S_k)P(A\backslash S_k)$$

Example: Total Probability

In a particular production run 20% of the chips are subjected to high levels of contamination, 30% to medium levels of contamination, and 50% to low levels of contamination.

Assume the following probabilities for product failure subject to levels of contamination

Level of Contamination	Prob. of Failure
High	0.10
Medium	0.01
Low	0.001



What is the probability that a product using one of these chips fails?

$$P(F)= P(H)\,P(F \backslash H)+ P(M)\,P(F \backslash M)+ P(L)\,P(F \backslash L)$$

$$=(0.20)(0.10)+(0.30)(0.01)+(0.50)(0.001)$$

$$=0.0235$$

BAYES' RULE

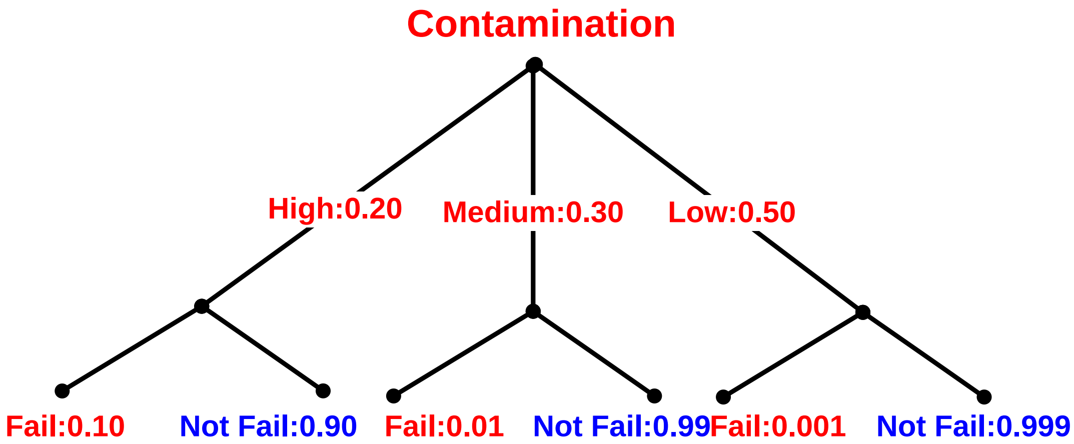
Let** S_1, S_2 ,…, S_k **represent k mutually exclusive and exhaustive subpopulations with prior probabilities** $P(S_1)$, $P(S_2)$,…,$P(S_k)$. **If an event A occurs, the posterior probability of** S_i **given A is the conditional probability,

$$P(S_i \setminus A) = \frac{P(S_i)P(A \setminus S_i)}{\sum_{i=1}^{k} P(S_i)P(A \setminus S_i)} \quad for\ \ i = 1,2,\ldots k.$$

$$\sum_{i=1}^{k} P(S_i) = 1$$

$$P(A) = \sum_{i=1}^{k} P(S_i)P(A \setminus S_i)$$

Example:



We have already calculated P(Fail)=P(F) using the Law of Total Probability. P(F)=0.0235

The conditional probability a high level of contamination was present when a failure occurred is to be determined.

$$P(High \setminus \text{F}) = \frac{P(High)P(F \setminus \text{High})}{P(F)} = \frac{(0.20)(0.10)}{0.0235} = 0.85106$$

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

(a) What is the probability that is defective?

(b) If a product were chosen randomly and found to be defective, what is the probability that it was made by machine B_3?

Solution

(a)

Let the events

A: the product is defective,

B_1: the product is made by machine B_1.

B_2: the product is made by machine B_2.

B_3: the product is made by machine B_3.

$P(A)=P(A\backslash B_1)P(B_1)+P(A\backslash B_2)P(B_2)+P(A\backslash B_3)P(B_3)$

$= (0.02)\,(0.3) + (0.03)\,(0.45) + (0.02)\,(0.25)$

$=0.0245.$

(b)

Using Bayes' rule

$P(B_3\backslash A)= P(A\backslash B_3)P(B_3)/P(A)=(0.02)(0.25)/0.0245$

$=0.005/0.0245=10/49=0.20408$

Example: Bayesian Network

Bayesian Networks are used on the Web sites of high-technology manufactures to allow customers to quickly diagnose problems with products.

An oversimplified example is presented here. A printer manufacturer obtained the following probabilities from a database of test results. **Printer failures** are associated with **three types of problems**:

- **Hardware with probability 0.1,**
- **Software with probability 0.6,**
- **Other (e.g. connecters) 0.3.**

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?

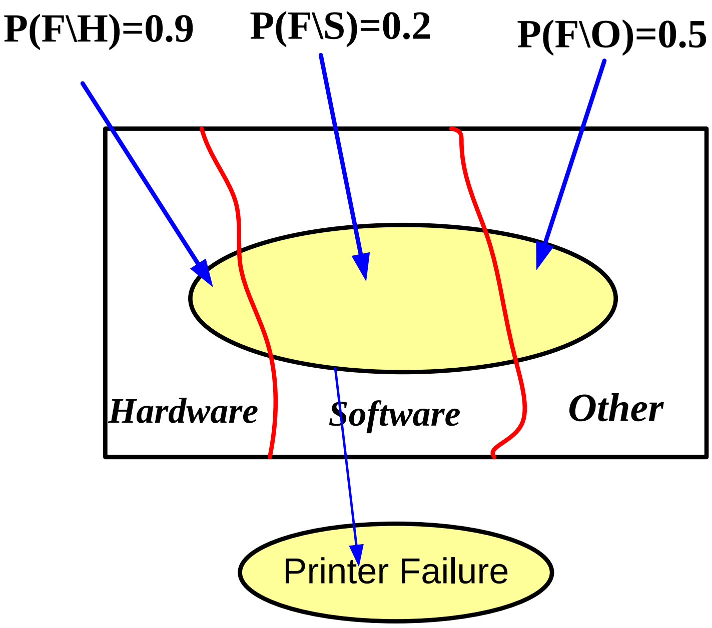
Let the events

H: Hardware problem

S: Software problem

O: Other problem and

F: printer failure.



<u>The most likely cause of the problem is the one that corresponds to the largest of</u>

P(H\F), P(S\F) and P(O\F).

In Bayes' Theorem the denominator is

P(F)=P(F\H)P(H)+P(F\S)P(S)+P(F\O)P(O)
= (0.9) (0.1) + (0.2) (0.6) + (0.5) (0.3) =0.36

Then

P(H\F) = P(F\H) P(H)/P(F) = (0.9)(0.1) /(0.36) =0.250

P(S\F) = P(F\S) P(S)/P(F) = (0.2)(0.6) /(0.36) =0.333

P(O\F) = P(F\O) P(O)/P(F) = (0.5)(0.3) /(0.36) =0.417

Notice that P (H\F) +P (S\F) +P (O\F) = 1 because one of the three types of problems is responsible for the failure. Because P (O\F) is largest, the most likely cause of the problem is in other category. A Web site dialog to diagnose the problem quickly should start with a check into that type of problem.