1.5

Bayes'Theorem and Its Applications

One of the important applications of the conditional probability is in the computation of unknown probabilities on the basis of the information supplied by the experiment or past records. For example, suppose an event has occurred through one of the various mutually exclusive events or reasons. Then the conditional probability that it has occurred due to a particular event or reason is called it as **inverse** or **posteriori probability**. These probabilities are computed by Bayes' theorem, named so after the British mathematician **Thomas Bayes** who propounded it in 1763. The revision of old (given) probabilities in the list of the additional information supplied by the experiment or past records is of extreme help in arriving at valid decisions in the face of uncertainty.

Bayes' Theorem (Rule for the Inverse Probability)

Let $E_1, E_2, \dots, E_n$ be n be mutually exclusive and exhaustive events in the sample space S with $P(E_i) \neq 0$ for $i = 1, 2, \dots, n$. Let A be an arbitrary event which is a subset of $\bigcup_{i=1}^{n} E_i$ such that $P(A) > 0$. Then

$$P(E_i \mid A) = \frac{P(E_i).P(A \mid E_i)}{\sum_{i=1}^{n} P(E_i).P(A \mid E_i)} = \frac{P(E_i).P(A \mid E_i)}{P(E_i)} \qquad \text{for } i = 1, 2, \dots, n$$

Proof: Since $A \subset \bigcup_{i=1}^{n} E_i$, we have $A = A \cap \left(\bigcup_{i=1}^{n} E_i \right) = \bigcup_{i=1}^{n} (A \cap E_i)$.

Since $(A \cap E_i) \subset E_i (i = 1, 2, \dots, n)$ are mutually exclusive events, we have by addition theorem of probability

$$P(A) = P\left(\bigcup_{i=1}^{n} (A \cap E_i) \right)$$

$$= \sum_{i=1}^{n} P(A \cap E_i)$$

$$\Rightarrow P(A)=\sum_{i=1}^{n} P(E_i).P(A|E_i)$$ (By multiplication theorem of probability)

Also we have

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

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

$$\Rightarrow P(E_i|A)=\frac{P(E_i).P(A|E_i)}{\sum_{i=1}^{n} P(E_i).P(A|E_i)} \quad \text{for } i = 1, 2, \dots, n$$

which is the Bayes' rule.

Note:

1. The probabilities $P(E_1), P(E_2), \dots, P(E_n)$ are known as the 'a priori probabilities', because they exist before we gain any information from the experiment itself.
2. The probabilities $P(A|\ E_i)\ i = 1, 2, \dots n$ are called 'likelihoods' because they indicate how likely the event A under consideration is to occur, given each and every a priori probability.
3. The probabilities $P(E_i|A), i = 1, 2, .., n$ are called 'posteriori probabilities' because they are determined after the results of the experiment are known.
4. $P(A)=\sum_{i=1}^{n} P(E_i).P(A|E_i)$ is known as **total probability**.
5. Bayes' theorem is extensively used by *business, management* and *engineering* executives in arriving at valid decisions in the face of uncertainty.

Example 1: In a bolt factory machines A, B, C manufacutre respectively 25%, 35% and 40% of the total. Of their output $5, 4, 2$ percent are known to be defective bolts. A bolt is drawn at random from the product and is found to be defective. What are the probabilities that it was manufactured by

(i) Machine A.

(ii) Machine B or C

Solution: Let E_1, E_2 and E_3 denote respectively the events that the bolt selected at random is manufactured by the machines A, B and C respectively and let E denote the event that it is defective. Then we have:

E_i	E_1	E_2	E_3	Total
$P(E_i)$	0.25	0.35	0.40	1
$P(E\|E_i)$	0.05	0.04	0.02	
$P(E \cap E_i) = P(E_i).P(E\|E_i)$	0.0125	0.0140	0.0080	$P(E) = 0.0345$

$$P(E) = \sum_{i=1}^{3} P(E_i).P(E|E_i) = 0.0345$$

(i) Hence, the probability that a defective bolt chosen at random is manufactured by factory A is given by Bayes' rule as:

$$P(E_1|E) = \frac{P(E_1)P(E|E_1)}{\sum P(E_i)P(E|E_i)} = \frac{0.0125}{0.0345} = 0.36$$

(ii) Similarly,

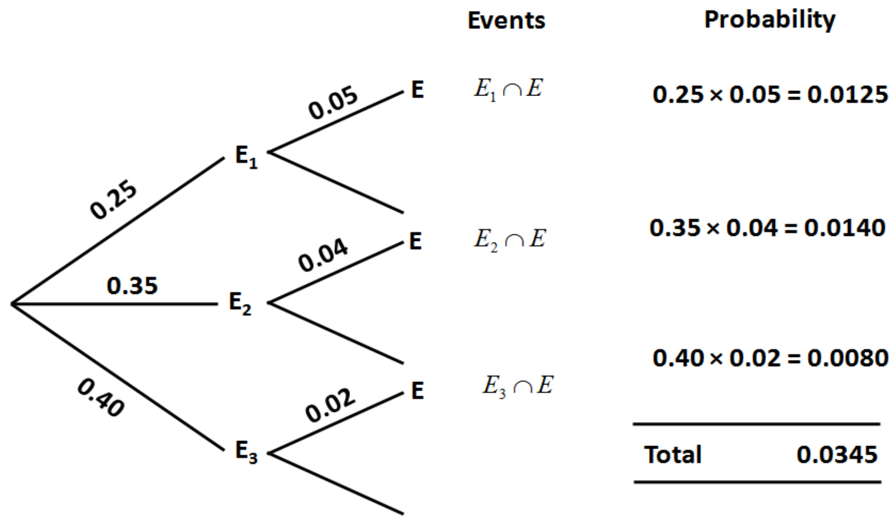
$$P(E_2|E) = \frac{0.0140}{0.0345} = \frac{28}{69} = 0.41\,; \qquad P(E_3|E) = \frac{0.0080}{0.0345} = \frac{16}{69} = 0.23$$

Hence, the probability that a defective bolt chosen at random is manufactured by machine B or C is:

$$P(E_2|E) + P(E_3|E) = 0.41 + 0.23 = 0.64$$

(OR Required probability is equal to $1 - P(E_1|E) = 1 - 0.36 = 0.64$)

Aliter: **TREE DIAGRAM**



From the above diagram the probability that a defective bolt is manufactured by factory A is

$$P(E_1|E) = \frac{0.0125}{0.0345} = 0.36$$

Similarly, $P(E_2|E) = \frac{0.0140}{0.0345} = 0.41$ and $P(E_3|E) = \frac{0.0080}{0.0345} = 0.23$

Hence, the probability that a defective bolt chosen at random is manufactured by machine B or C is:

$$P(E_2|E) + P(E_3|E) = 0.41 + 0.23 = 0.64$$

(OR Required probability is equal to $1 - P(E_1|E) = 1 - 0.36 = 0.64$)

Remark: Since $P(E_3)$ is greatest, on the basis of 'a *priori*' probabilities alone, we are likely to conclude that a defective bolt drawn at random from the product is manufactured by machine C. After using the additional information we obtained the *posterior* probabilities which give $P(E_2|E)$ as maximum. Thus, we shall now say that it is probable that the defective bolt has been manufactured by machine B, a result which is different from the earlier conclusion. However, latter conclusion is a much valid conclusion as it is based on the entire information at our disposal. Thus, Bayes's rule provides a very powerful tool in improving the quality of probability and this helps the management executives in arriving at valid decisions in the face of uncertainty. Thus, the additional information reduces the importance of the prior probabilities. The only requirement for the use of *Bayesian rule* is that all the hypotheses under consideration must be valid and that none is assigned 'a prior' probability 0 or 1.

Example 2: In a railway reservation office, two clerks are engaged in checking reservation forms. On an average, the first clerk checks 55% of the forms, while the second does the remaining. The first clerk has an error rate of 0.03 and second has an error rate of 0.02 , A reservation form is selected at random from the total number of forms checked during a day, and is found to have an error. Find the probability that it was checked (i) by the first (ii) by the second clerk.

Solution: Let us define the following events:

E_1: The selected form is checked by clerk 1.

E_2: The selected form is checked by clerk 2.

E: The selected form has an error.

Then we are given:

$$P(E_1) = 55\% = 0.55 \;; \quad P(E_2) = 45\% = 0.45 \;;$$

$$P(E|E_1) = 0.03 \quad ; \quad P(E|E_2) = 0.02$$

Required to find $P(E_1|E)$ and $P(E_2|E)$. By Bayes' Rule the probability that the form containing the error was checked by clerk 1, is given by;

$$P(E_1|E) = \frac{P(E_1)\,P(E_1|E)}{P(E_1)\,P(E|E_1)+P(E_2)P(E|E_2)} = \frac{0.55\times 0.03}{0.55\times 0.03+0.45\times 0.02}$$

$$= \frac{0.0165}{0.0165+0.0090} = \frac{0.0165}{0.0255} = 0.647$$

Similarly, the probability that the form containing the error was checked by clerk 2, is given by

$$P(E_2|E) = \frac{P(E_2)\,P(E|E_2)}{P(E_1)P(E|E_1)+P(E_2)P(E|E_2)} = \frac{0.45\times 0.02}{0.55\times 0.03+0.45\times 0.02} = \frac{0.0090}{0.0255} = 0.353$$

(OR $P(E_2|E) = 1 - P(E_1|E) = 1 - 0.647 = 0.353$)

Example 3: The results of an investigating by an expert on a fire accident in a skyscraper are summarized below:

(i) **Prob. (there could have been short circuit)** $= \mathbf{0.8}$

(ii) **Prob. (LPG cylinder explosion)** $= \mathbf{0.2}$

(iii) **Chance of fire accident is** 30% **given a short circuit and** 95% **given an LPG explosion.**

Based on these, what do you think is the most probable cause of fire?

Solution: Let us define the following events:

E_1: Short circuit ; E_2: LPG explosion ; E: Fire accident

Then, we are given:

$$P(E_1) = 0.8 \; ; \quad P(E_2) = 0.2 \; ;$$

$$P(E|E_1) = 0.30 \; ; \quad P(E|E_2) = 0.95$$

By Bayes' Rule:

$$P(E_1|E) = \frac{P(E_1).P(E|E_1)}{P(E_1)P(E|E_1)+P(E_2)P(E|E_2)} = \frac{0.80\times0.30}{0.8\times0.30+0.2\times0.95} = \frac{0.240}{0.240+0.190} = \frac{24}{43}$$

$$P(E_2|E) = \frac{P(E_2)P(E|E_2)}{P(E_1)P(E|E_1)+P(E_2)P(E|E_2)} = \frac{0.190}{0.430} = \frac{19}{43}$$

(OR $P(E_2|E) = 1 - P(E_1|E) = 1 - \frac{24}{43} = \frac{19}{43}$)

Since $P(E_1|E) > P(E_2|E)$, short circuit is the most probable cause of fire.

Example 4: The contents of urns I, II and III are respectively as follows:

1 white, 2 black and 3 red balls,

2 white, 1 black and 1 red balls, and

4 white, 5 black and 3 red balls.

One urn is chosen at random and two balls drawn. They happen to be white and red. What is the probability that they came from urns I, II, III?

Solution:

Let E_1, E_2 and E_3 denote the events of choosing 1^{st} , 2^{nd} and 3^{rd} urn respectively and let E be the event that the two balls drawn from the selected urn are white and red. Then we have:

	E_1	E_2	E_3
$P(E_i)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
$P(E\|E_i)$	$\frac{1\times 3}{6_{C_2}} = \frac{1}{5}$	$\frac{2\times 1}{4_{C_2}} = \frac{1}{3}$	$\frac{4\times 3}{12_{C_2}} = \frac{2}{11}$
$P(E\cap E_i) = P(E_i)\times P(E\|E_i)$	$\frac{1}{3}\times\frac{1}{5} = \frac{1}{15}$	$\frac{1}{3}\times\frac{1}{3} = \frac{1}{9}$	$\frac{1}{3}\times\frac{2}{11} = \frac{2}{33}$

We have:

$$\sum P(E_i)\, P(E|E_i) = \frac{1}{15} + \frac{1}{9} + \frac{2}{33} = \frac{33+55+30}{495} = \frac{118}{495}$$

Hence by Bayes's rule, the probability that the two white and red balls drawn are from 1^{st} urn is:

$$P(E_1|E) = \frac{P(E_1)P(E|E_1)}{\sum P(E_i)P(E|E_i)} = \frac{\frac{1}{15}}{\frac{118}{495}} = \frac{33}{118}$$

Similarly, we have

$$P(E_2|E) = \frac{P(E_2)P(E|E_2)}{\sum P(E_i)P(E|E_i)} = \frac{\frac{1}{9}}{\frac{118}{495}} = \frac{55}{118}$$

and $P(E_3|E) = \frac{\frac{2}{33}}{\frac{118}{495}} = \frac{30}{118}$ (Or $P(E_3|E) = 1 - \frac{33}{118} - \frac{55}{118} = \frac{30}{118}$)