MODULE 1

PROBABILITY

LECTURE 5

Topics

1.3.2 Bayes' Theorem

1.3.2 Bayes' Theorem

The following theorem provides a method for finding the probability of occurrence of an event in a past trial based on information on occurrences in future trials.

***1.3.2* Theorem 3.4 (*Bayes' Theorem*)**

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $\{E_i : i \in \Lambda\}$ be a countable collection of mutually exclusive and exhaustive events with $P(E_i) > 0, i \in \Lambda$. Then, for any event $E \in \mathcal{F}$ with $P(E) > 0$, we have

$$P(E_j|E) = \frac{P(E|E_j)P(E_j)}{\sum_{i \in \Lambda} P(E|E_i)P(E_i)}, \qquad j \in \Lambda \cdot$$

Proof. We have, for $j \in \Lambda$,

$$P(E_j|E) = \frac{P(E_j \cap E)}{P(E)}$$

$$= \frac{P(E|E_j)P(E_j)}{P(E)}$$

$$= \frac{P(E|E_j)P(E_j)}{\sum_{i \in \Lambda} P(E|E_i)P(E_i)} \text{ (using Theorem of Total Probability).} \blacksquare$$

Remark 3.2

(i) Suppose that the occurrence of any one of the mutually exclusive and exhaustive events $E_i, i \in \Lambda$, causes the occurrence of an event E. Given that the event E has occurred, Bayes' theorem provides the conditional probability that the event E is caused by occurrence of event $E_j, j \in \Lambda$.

(ii) In Bayes' theorem the probabilities $P(E_j), j \in \Lambda$, are referred to as *prior probabilities* and the probabilities $P(E_j|E), j \in \Lambda$, are referred to as *posterior probabilities*. ■

To see an application of Bayes' theorem let us revisit Example 3.4.

Example 3.5

Urn U_1 contains 4 white and 6 black balls and urn U_2 contains 6 white and 4 black balls. A fair die is cast and urn U_1 is selected if the upper face of die shows five or six dots. Otherwise urn U_2 is selected. A ball is drawn at random from the selected urn.

(i) Given that the drawn ball is white, find the conditional probability that it came from urn U_1;

(ii) Given that the drawn ball is white, find the conditional probability that it came from urn U_2.

Solution. Define the events:

W : drawn ball is white;

$$\left.\begin{array}{l} E_1 : \text{ urn } U_1 \text{ is selected} \\ E_2 : \text{urn } U_2 \text{ is selected} \end{array}\right\} \text{mutually exclusive \& exhaustive events}$$

(i) We have

$$P(E_1|W) = \frac{P(W|E_1)P(E_1)}{P(W|E_1)P(E_1) + P(W|E_2)P(E_2)}$$

$$= \frac{\frac{4}{10} \times \frac{2}{6}}{\frac{4}{10} \times \frac{2}{6} + \frac{6}{10} \times \frac{4}{6}}$$

$$= \frac{1}{4}.$$

(ii) Since E_1 and E_2 are mutually exclusive and $P(E_1 \cup E_2|W) = P(\Omega|W) = 1$,we have

$$P(E_2|W) = 1 - P(E_1|W)$$

$$= \frac{3}{4}. \blacksquare$$

In the above example

$$P(E_1|W) = \frac{1}{4} < \frac{1}{3} = P(E_1),$$

and $$P(E_2|W) = \frac{3}{4} > \frac{2}{3} = P(E_2),$$

i.e.,

(i) the probability of occurrence of event E_1 decreases in the presence of the information that the outcome will be an element of W;

(ii) the probability of occurrence of event E_2 increases in the presence of information that the outcome will be an element of W.

These phenomena are related to the concept of association defined in the sequel.

Note that

$$P(E_1|W) < P(E_1) \Leftrightarrow P(E_1 \cap W) < P(E_1)P(W),$$

and

$$P(E_2|W) > P(E_2) \Leftrightarrow P(E_2 \cap W) > P(E_2)P(W).$$

Definition 3.2

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let A and B be two events. Events A and B are said to be

(i) *negatively associated* if $P(A \cap B) < P(A)P(B)$;

(ii) *positively associated* if $P(A \cap B) > P(A)P(B)$;

(iii) *independent* if $P(A \cap B) = P(A)P(B)$. ■

Remark 3.3

(i) If $P(B) = 0$ then $P(A \cap B) = 0 = P(A)P(B), \forall A \in \mathcal{F}$, i.e., if $P(B) = 0$ then any event $A \in \mathcal{F}$ and B are independent;

(ii) If $P(B) > 0$ then A and B are independent if, and only if, $P(A|B) = P(A)$, i.e., if $P(B) > 0$, then events A and B are independent if, and only if, the availability of the information that event B has occurred does not alter the probability of occurrence of event A. ■

Now we define the concept of independence for arbitrary collection of events.

Definition 3.3

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $\Lambda \subseteq \mathbb{R}$ be an index set and let $\{E_\alpha : \alpha \in \Lambda\}$ be a collection of events in $\mathcal{F}$.

(i) Events $\{E_\alpha : \alpha \in \Lambda\}$ are said to be *pairwise independent* if any pair of events E_α and $E_\beta, \alpha \neq \beta$ in the collection $\{E_j : j \in \Lambda\}$ are independent. i.e., if $P(E_\alpha \cap E_\beta) = P(E_\alpha)P(E_\beta)$, whenever $\alpha, \beta \in \Lambda$ and $\alpha \neq \beta$;

(ii) Let $\Lambda = \{1, 2, \dots, \mathrm{n}\}$, for some $n \in \mathbb{N}$, so that $\{E_\alpha : \alpha \in \Lambda\} = \{E_1, \dots, E_n\}$ is a finite collection of events in $\mathcal{F}$. Events $E_1, \dots, E_n$ are said to be *independent* if, for any sub collection $\{E_{\alpha_1}, \dots, E_{\alpha_k}\}$ of $\{E_1, \dots, E_n\}$ $(k = 2,3, \dots, n)$

$$P\left(\bigcap_{j=1}^{k} E_{\alpha_j}\right) = \prod_{j=1}^{k} P\left(E_{\alpha_j}\right). \tag{3.6}$$

(iii) Let $\Lambda \subseteq \mathbb{R}$ be an arbitrary index set. Events $\{E_\alpha : \alpha \in \Lambda\}$ are said to be independent if any finite sub collection of events in $\{E_\alpha : \alpha \in \Lambda\}$ forms a collection of independent events. ■

Remark 3.4

(i) To verify that n events $E_1, \dots, E_n \in \mathcal{F}$ are independent one must verify $2^n - n - 1 \left(= \sum_{j=2}^{n} \binom{n}{j}\right)$ conditions in (3.6). For example, to conclude that three events E_1, E_2 and E_3 are independent, the following 4 $(= 2^3 - 3 - 1)$ conditions must be verified:

$$P(E_1 \cap E_2) = P(E_1)P(E_2);$$
$$P(E_1 \cap E_3) = P(E_1)P(E_3);$$
$$P(E_2 \cap E_3) = P(E_2)P(E_3);$$
$$P(E_1 \cap E_2 \cap E_3) = P(E_1)P(E_2)P(E_3).$$

(ii) If events $E_1, \dots, E_n$ are independent then, for any permutation $(\alpha_1, \dots, \alpha_n)$ of $(1, \dots, n)$, the events $E_{\alpha_1}, \dots, E_{\alpha_n}$ are also independent. Thus the notion of independence is symmetric in the events involved.

(iv) Events in any subcollection of independent events are independent. In particular independence of a collection of events implies their pairwise independence. ■

The following example illustrates that, in general, pairwise independence of a collection of events may not imply their independence.

Example 3.6

Let $\Omega = \{1, 2, 3, 4\}$ and let $\mathcal{F} = \mathcal{P}(\Omega)$, the power set of Ω. Consider the probability space $(\Omega, \mathcal{F}, \mathrm{P})$, where $P(\{i\}) = \frac{1}{4}, i = 1, 2, 3, 4$. Let $A = \{1, 4\}$, $B = \{2, 4\}$ and $C = \{3, 4\}$. Then,

$$P(A) = P(B) = P(C) = \frac{1}{2},$$

$$P(A \cap B) = P(A \cap C) = P(B \cap C) = P(\{4\}) = \frac{1}{4},$$

and $$P(A \cap B \cap C) = P(\{4\}) = \frac{1}{4} \cdot$$

Clearly,

$P(A \cap B) = P(A)P(B); P(A \cap C) = P(A)P(C)$, and $P(B \cap C) = P(B)P(C)$,

i.e., A, B and C are pairwise independent.

However,

$$P(A \cap B \cap C) = \frac{1}{4} \neq P(A)P(B)P(C).$$

Thus A, B and C are not independent. ■