Conditional Probability and Bayes Formula

Consider a random experiment with sample space $\mathcal{S}$. In many situations we may not be interested in the whole sample space but a subset of the sample space.

For example, suppose that we toss two fair dice. Then $\mathcal{S} = \{(i,j) : i,j \in \{1,2,3,4,5,6\}\}$ and $P((i,j)) = \frac{1}{36}$. Suppose we observe that the first die is four. If the first die is a four, then there can be at most six possible outcomes, namely, $(4,1)$, $(4,2)$, $(4,3)$, $(4,4)$, $(4,5)$, $(4,6)$. So given that the first die is a four, the probability that the sum of the two dice equals six is $1/6$. Let E and F be the event that the sum of the dice is six and the event that the first die is a four respectively, then the probability obtained is called the conditional probability that E occurs given that F has occurred.

Definition 1. *Let $(\mathcal{S}, \Sigma, P)$ be a probability space and F be the fixed event with $P(F) > 0$. Then the probability that an event E occurs given that F has occurred is called the conditional probability of event E given that the outcomes of the experiment is in F or simply the conditional probability of E given F. It is denoted by $P(E|F)$ and is defined by*

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

Theorem 2. *Let $(\mathcal{S}, \Sigma, P)$ be a probability space and F be the fixed event with $P(F) > 0$. Then $(\mathcal{S}, \Sigma, P_F)$, where $P_F(E) = P(E|F)$ for all $E \in \Sigma$, is a probability space.*

Proof. Exercise □

Example 3. *Suppose cards numbered one to ten are placed in a hat, mixed up, and then one of the card is drawn. If we are told that the number on the drawn card is at least five, then what is the conditional probability that it is ten?*

Solution: Let E be the event that the number on the drawn card is ten and F be the event that the number on the drawn card is at least five.

Then $P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{1/10}{6/10} = 1/6$.

Example 4. *Suppose that each of three men at a party throws his hat into the center of the room. The hats are first mixed up and then each man randomly select a hat. What is the probability that none of the three men select his own hat?*

Solution: For $i = 1,2,3$, let E_i be the event that the ith man selects his own hat. Then $P(E_i) = 1/3$, for each $i = 1,2,3$.

Given that the i-th man has selected his own hat, then there remain two hats that the j-th man may select, and as one of these two is his own hat, so $P(E_j|E_i) = 1/2$ for $i \neq j$. Therefore, $P(E_i \cap E_j) = P(E_i)P(E_j|E_i) = 1/6$ for $i \neq j$. Also $P(E_1 \cap E_2 \cap E_3) = P(E_1 \cap E_2)P(E_3|E_1 \cap E_2) = \frac{1}{6}P(E_3|E_1 \cap E_2)$.

However, given that the first two men get their own hats, it follows that the third man must also get his own hat (since there are no other hats left). That is, $P(E_3|E_1 \cap E_2) = 1$. So $P(E_1 \cap E_2 \cap E_3) = \frac{1}{6}$

Now, the probability that at least one of them selects his own hat is, $P(E_1 \cup E_2 \cup E_3) = P(E_1) + P(E_2) + P(E_3) - P(E_1 \cap E_2) - P(E_2 \cap E_3) - P(E_1 \cap E_3) + P(E_1 \cap E_2 \cap E_3) = 2/3$.

Hence the probability that none of the three men select his own hat is $1 - P(E_1 \cup E_2 \cup E_3) = 1/3$.

Example 5. *Suppose an urn contains seven black balls and five white balls. We draw two balls from the urn without replacement. Assuming that each ball in the urn is equally likely to be drawn, what is the probability that both drawn balls are black?*

Solution: Let E be the event that the first drawn ball is black and F be the event that the second drawn ball is black.

Then $P(E) = 7/12$ and $P(F|E) = 6/11$. Therefore, $P(E \cap F) = P(E)P(F|E) = 42/132$.

Independent Events: Two events E and F are said to be independent if $P(E \cap F) = P(E)P(F)$. Two events E and F are said to be dependent if E and F are not independent.

Remark 6. (1) *Suppose $P(F) > 0$. Then $P(E|F) = \frac{P(E\cap F)}{P(F)}$. Assume E and F are independent i.e., $P(E \cap F) = P(E)P(F)$. Then $P(E|F) = P(E)$. This implies that if $P(F) > 0$, then E and F are independent if and only if $P(E|F) = P(E)$.*

In other words, E and F are independent if and only if the availability of the information that event F has occurred does not alter the probability of occurrence of event E.

(2) *If $P(F) = 0$, then $P(E \cap F) = 0 = P(E)P(F)$, for every event E, that is, if $P(F) = 0$, then any event E and F are independent.*

Definition 7. (1) *Events $\{E_i : i \in \wedge\}$, where $\wedge$ is an index set, are said to be pairwise independent if any pair of events E_i and E_j, $i \neq j$, are independent, i.e., $P(E_i \cap E_j) = P(E_i)P(E_j)$ $i \neq j$.*

(2) *Events $E_1, E_2, \ldots, E_n$ are said to be independent if for any sub-collection $\{E_{i_1}, E_{i_2}, \ldots, E_{i_k}\}$, $2 \leq k \leq n$, we have $P(\cap_{j=1}^{k} E_{i_j}) = \prod_{j=1}^{k} P(E_{i_j})$.*

By definition, it is clear that independence of a finite collection of events implies pairwise independence of events but the converse need not be true. Also to verify that n events $E_1, E_2, \cdots, E_n$ are independent, we have to verify $2^n - n - 1$ conditions.

The following example shows that pairwise independence does not imply independence.

Example 8. *Let a sample space $\mathcal{S} = \{1, 2, 3, 4\}$ with $P(i) = 1/4$, for each $i = 1, 2, 3, 4$. Consider the following events: $A = \{1, 4\}$, $B = \{2, 4\}$ and $C = \{3, 4\}$. Then $P(A) = P(B) = P(C) = 1/2$, $P(A \cap B) = P(B \cap C) = P(C \cap A) = 1/4$, and $P(A \cap B \cap C) = 1/4$. Clearly A, B and C are pairwise independent but not independent.*

Example 9. *Suppose we toss two fair dice. Let E_1 be the event that the sum of the dice is six, E_2 be the event that the sum of the dice is seven, and F be the event that the first die equals four. Then*

$$P(E_1) = 5/36, P(E_2) = 1/6, P(F) = 1/6 \text{ and } P(E_1 \cap F) = P(E_2 \cap F) = 1/36.$$

Since $P(E_1 \cap F) \neq P(E_1)P(F)$, E_1 and F are not independent. On the other hand, since $P(E_2 \cap F) = P(E_2)P(F)$, E_2 and F are independent.

Intuitively it is clear, because chance of getting a total of six depends on the outcome of the first die and hence E_1 and F cannot be independent while for every outcome of the first die, we have a chance of getting a total of seven and hence E_2 and F are independent.

Bayes Formula

Let E and F be two events. Then we have

$$E = (E \cap F) \cup (E \cap F^c)$$

Since $E \cap F$ and $E \cap F^c$ are mutually exclusive,

$$P(E) = P(E \cap F) + P(E \cap F^c)$$

Suppose $0 < P(F) < 1$. Then

$$P(E) = P(E|F)P(F) + P(E|F^c)P(F^c)$$

Now we will generalize this as follows:

Suppose that $F_1, F_2, \cdots, F_n$ are mutually exclusive and exhaustive events, that is, $F_i \cap F_j = \emptyset$ for $i \neq j$ and $P(\cup_{i=1}^n F_i) = 1$.

Let E be any event and $F = \cup_{i=1}^n F_i$. Since $P(F) = 1$, $P(F^c) = 0$. Then

$$\begin{aligned} P(E) &= P(E \cap F) \\ &= P(\cup_{i=1}^n (E \cap F_i)) \\ &= \sum_{i=1}^n P(E \cap F_i) \end{aligned}$$

Suppose $P(F_i) > 0$, for all $1 \leq i \leq n$. Then we have

$$\boxed{P(E) = \sum_{i=1}^n P(E|F_i)P(F_i)}$$

This formula is called the **total probability rule**.

Now, for $P(E) > 0$, we have

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

$$\boxed{P(F_j|E) = \frac{P(E|F_j)P(F_j)}{\sum_{i=1}^n P(E|F_i)P(F_i)}}$$

This is called the **Bayes formula**.

Example 10. *Consider two urns. The first contains two white and seven black ball, and the second contains five white and six black balls. We flip a fair coin and then draw ball from the first urn or the second urn depending on whether the outcome was heads or tails. What is the conditional probability that the outcome of the toss was heads given that a white ball was selected?*

Solution: Let W be the event that a white ball is drawn and H be the event that the coin comes up head.

Then $P(H|W) = \frac{P(H \cap W)}{P(W)} = \frac{P(W|H)P(H)}{P(W)} = \frac{P(W|H)P(H)}{P(W|H)P(H)+P(W|H^c)P(H^c)}$

Hence $P(H|W) = \frac{2/9 \times 1/2}{(2/9 \times 1/2)+(5/11 \times 1/2)} = \frac{22}{67}$.

Example 11. *Urn 1 contains one white and two black marbles, Urn 2 contains one black and two white marbles, and Urn 3 contains three black and three white marbles. A die is rolled. If a* $1, 2$ *or* 3 *shows up, Urn 1 is selected, if a* 4 *shows up, Urn 2 is selected, and if a* 5 *or* 6 *shows up, Urn 3 is selected. A marble is then drawn at random from the urn selected. Let* A *be the event that the drawn marble is white and* U, V, W *respectively denotes the events that the urn selected is* $1, 2, 3$. *What is the probability that urn 2 is selected given that the marble drawn is white?*

Solution: $P(V|A) = \frac{P(A\cap V)}{P(A)} = \frac{P(A|V)P(V)}{P(A|U)P(U)+P(A|V)P(V)+P(A|W)P(W)}$

Hence $P(V|A) = \frac{1/6\times 2/3}{(3/6\times 1/3)+(1/6\times 2/3)+(2/6\times 3/6)} = \frac{1}{4}$.

Example 12. *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.*

(1) *Given that the drawn ball is white, find the conditional probability that it came from Urn* U_1.
(2) *Given that the drawn ball is white, find the conditional probability that it came from Urn* U_2.

Solution: Let W be the event that the drawn ball is white, E_1 be the event that the Urn U_1 is selected and E_2 be the event that the Urn U_2 is selected. Clearly E_1 and E_2 are mutually exclusive and exhaustive events.

(1) $P(E_1|W) = \frac{P(E_1\cap W)}{P(W)}) = \frac{P(W|E_1)P(E_1)}{P(W|E_1)P(E_1)+P(W|E_2)P(E_2)} = \frac{4/10\times 2/6}{(4/10\times 2/6)+(6/10\times 4/6)} = \frac{1}{4}$
(2) $P(E_2|W) = \frac{P(E_2\cap W)}{P(W)}) = \frac{P(W|E_2)P(E_2)}{P(W|E_1)P(E_1)+P(W|E_2)P(E_2)} = \frac{6/10\times 4/6}{(4/10\times 2/6)+(6/10\times 4/6)} = \frac{3}{4}$