

Conditional Probability and Independence

1.1. Conditional Probability

Knowledge that a particular event A has occurred will change our assessment of the probabilities of other event B . In such an example, the terminology “conditional probability” is used.

An experiment is conducted with sample space Ω , given event B has occurred. The probability event A occurs given event B has occurred, written $P(A|B)$, is the probability of A relative to the reduced sample space B . ($A \cap B$) is the subset of B for which A is true, so the probability of A given B should be proportional to $P(A \cap B)$, say $P(A|B) = kP(A \cap B)$. Similarly, $P(A^c|B) = kP(A^c \cap B)$. Thus,

$$P(A|B) + P(A^c|B) = k [P(A \cap B) + P(A^c \cap B)] = kP [(A \cap B) \cup (A^c \cap B)] = kP(B) = 1$$

Therefore, $k = \frac{1}{P(B)}$. And $P(A|B) = kP(A \cap B) = \frac{P(A \cap B)}{P(B)}$.

DEFINITION 1. Suppose that A and B are events defined on some sample space Ω . If $P(B) > 0$ then $P(A|B) = \frac{P(A \cap B)}{P(B)}$ is called the conditional probability of A given B .

EXAMPLE 2. A coin is flipped twice. Assuming that all four points in the sample space $S = (h, h), (h, t), (t, h), (t, t)$ are equally likely, what is the conditional probability that both flips land on heads, given that (a) the first flip lands on heads? (b) at least one flip lands on heads?

Solution. Let $B = (h, h)$ be the event that both flips land on heads; let $F = (h, h), (h, t)$ be the event that the first flip lands on heads; and let $A = (h, h), (h, t), (t, h)$ be the event that at least one flip lands on heads. The probability for (a) can be obtained from

$$P(B|F) = \frac{P(BF)}{P(F)} = \dots$$

For (b),

$$P(B|A) = \frac{P(BA)}{P(A)} = \dots$$

From definition 1, $P(A \cap B) = P(B)P(A|B)$. A generalization of the probability of the intersection of an arbitrary number of events, is sometimes referred to as **the multiplication rule**.

$$P(E_1 E_2 E_3 \dots E_n) = P(E_1)P(E_2|E_1)P(E_3|E_1 E_2) \dots P(E_n|E_1 \dots E_{n-1})$$

EXAMPLE 3. An ordinary deck of 52 playing cards is randomly divided into 4 piles of 13 cards each. Compute the probability that each pile has exactly 1 ace.

Solution.

- (1) First way to solve this problem is define events $E_i, i = 1, 2, 3, 4$, be the events that the i th pile has exactly one ace.

$$P(E_1) = \frac{\binom{4}{1} \binom{48}{12}}{\binom{52}{13}}$$

$$P(E_2|E_1) = \frac{\binom{3}{1} \binom{36}{12}}{\binom{39}{13}}$$

$$P(E_3|E_1E_2) = \frac{\binom{2}{1} \binom{24}{12}}{\binom{26}{13}}$$

$$P(E_4|E_1E_2E_3) = \frac{\binom{1}{1} \binom{12}{12}}{\binom{13}{13}} = 1$$

The desired probability is $P(E_1E_2E_3E_4)$, and by the multiplication rule

$$P(E_1E_2E_3E_4) = P(E_1)P(E_2|E_1)P(E_3|E_1E_2)P(E_4|E_1E_2E_3) = \dots$$

- (2) Another way to solve this problem is by defining events $E_i, i = 1, 2, 3, 4$, as follows $E_1 = \{\text{the ace of spades is in any one of the piles}\}$, $E_2 = \{\text{the ace of spades and the ace of hearts are in different piles}\}$, $E_3 = \{\text{the aces of spades, hearts, and diamonds are all in different piles}\}$, $E_4 = \{\text{all 4 aces are in different piles}\}$. The desired probability is $P(E_1E_2E_3E_4)$, and by the multiplication rule,

$$P(E_1E_2E_3E_4) = P(E_1)P(E_2|E_1)P(E_3|E_1E_2)P(E_4|E_1E_2E_3)$$

Now, $P(E_1) = 1$ since E_1 is the sample space. Also, $P(E_2|E_1) = \frac{39}{51}$ since the pile containing the ace of spades will receive 12 of the remaining 51 cards, and $P(E_3|E_1E_2) = \frac{26}{50}$ since the piles containing the aces of spades and hearts will receive 24 of the remaining 50 cards. Finally, $P(E_4|E_1E_2E_3) = \frac{13}{49}$. Therefore, the probability that each pile has exactly 1 ace is $P(E_1E_2E_3E_4) = \frac{39 \cdot 26 \cdot 13}{51 \cdot 50 \cdot 49} \approx 0.105$. That is, there is approximately a 10.5 percent chance that each pile will contain an ace.

PROPOSITION 4. (**Law of total probability**) If B_1, B_2, \dots are disjoint events with $P(B_k) > 0$ for all k and $\bigcup_{k=1}^{\infty} B_k = \Omega$ then

$$P(A) = \sum_{k=1}^{\infty} P(B_k) P(A|B_k)$$

PROOF. Since $A = A \cap \left(\bigcup_{k=1}^{\infty} B_k \right)$ and $A \cap B_1, A \cap B_2, \dots$ are disjoint, then

$$\begin{aligned} P(A) &= P\left(A \cap \left(\bigcup_{k=1}^{\infty} B_k\right)\right) \\ &= P\left(\bigcup_{k=1}^{\infty} (A \cap B_k)\right) \\ &= \sum_{k=1}^{\infty} P(A \cap B_k) \\ &= \sum_{k=1}^{\infty} P(B_k) P(A|B_k) \end{aligned}$$

□

Let A and B be any events and $B \cup B^c = \Omega$. Then from proposition 4,

$$P(A) = P(B)P(A|B) + P(B^c)P(A|B^c)$$

This is an extremely useful formula, because its use often enables us to determine the probability of an event by first “conditioning” upon whether or not some second event has occurred.

EXAMPLE 5. An insurance company believes that people can be divided into two classes, those who are accident prone and those who are not. The company’s statistics show that an accident-prone person will have an accident at some time within a fixed 1-year period with probability .4, whereas this probability decreases to .2 for a person who is not accident prone. If 30 percent of the population is assumed accident prone, what is the probability that a new policyholder will have an accident within a year of purchasing a policy?

Solution. Let A_1 denote the event that the policyholder will have an accident within a year of purchasing the policy, and let A denote the event that the policyholder is accident prone. Hence, the desired probability is given by

$$P(A_1) = P(A)P(A_1|A) + P(A^c)P(A_1|A^c) = (0.3)(0.4) + (0.7)(0.2) = 0.26$$

A simple corollary of the law of total probability is Bayes’ Theorem.

PROPOSITION 6. (**Bayes’ Theorem**) Suppose that B_1, B_2, \dots are disjoint sets with $P(B_k) > 0$ for all k and $\bigcup_{k=1}^{\infty} B_k = \Omega$. Then for any event A ,

$$P(B_j|A) = \frac{P(B_j)P(A|B_j)}{\sum_{k=1}^{\infty} P(B_k)P(A|B_k)}$$

PROOF. By definition 1, $P(B_j|A) = \frac{P(AB_j)}{P(A)}$ and $P(AB_j) = P(B_j)P(A|B_j)$. The conclusion follows by applying the law of total probability to $P(A)$. □

EXAMPLE 7. In answering a question on a multiple-choice test, a student either knows the answer or guesses. Let p be the probability that the student knows the answer and $1-p$ be the probability that the student guesses. Assume that a student who guesses at the answer will be correct with probability $\frac{1}{m}$, where m is the number of multiple-choice alternatives. What is the conditional probability that a student knew the answer to a question given that he or she answered it correctly?

Solution. Let C and K denote, respectively, the events that the student answers the question correctly and the event that he or she actually knows the answer. Now,

$$P(K|C) = \frac{P(K)P(C|K)}{P(K)P(C|K) + P(K^c)P(C|K^c)} = \frac{p \cdot 1}{p \cdot 1 + (1-p) \cdot \frac{1}{m}} = \frac{mp}{1 + (m-1)p}$$

For example, if $m = 5$, $p = \frac{1}{2}$, then the probability that the student knew the answer to a question he or she answered correctly is $\frac{5}{6}$.

1.2. Independence

A card is selected at random from an ordinary deck of 52 playing cards. If E is the event that the selected card is an ace and F is the event that it is a spade, then E and F are independent.

Two coins are flipped, and all 4 outcomes are assumed to be equally likely. If A is the event that the first coin lands on heads and B the event that the second lands on tails, then A and B are independent.

DEFINITION 8. Events A and B are said to be independent if $P(A \cap B) = P(A)P(B)$.

Consider an experiment where two fair dice are tossed. Let E denote the event that the sum of the dice is 6 and F_1 denote the event that the first die equals 4. Then $P(E \cap F_1) = P((4, 2)) = \frac{1}{36}$ whereas $P(E)P(F_1) = \frac{5}{36} \cdot \frac{1}{6} = \frac{5}{216}$. Hence, E and F_1 are not independent.

Let D be the event that the sum of the dice equals 7. D and F_1 are independent. Why?

PROPOSITION 9. If A and B are independent, then A and B^c are independent, A^c and B are independent, and A^c and B^c are independent

PROOF. Assume that A and B are independent. Since $A = AB \cup AB^c$ and AB and AB^c are obviously mutually exclusive, then

$$P(A) = P\left(AB \cup AB^c\right) = P(AB) + P(AB^c)$$

or, equivalently,

$$P(AB^c) = P(A) - P(AB) = P(A) - P(A)P(B) = P(A)(1 - P(B)) = P(A)P(B^c)$$

and the result is proved. \square

Notion of independence can be extended to a finite or countably infinite collection of events.

DEFINITION 10. A_1, A_2, \dots, A_n (or A_1, A_2, \dots) are (mutually) independent events if for any finite subcollection $A_{i_1}, A_{i_2}, \dots, A_{i_k}$,

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = \prod_{j=1}^k P(A_{i_j})$$

Thus if A_1, A_2, \dots, A_n are independent then A_i is independent of A_j for $i \neq j$; that is, **mutual independence** implies **pairwise independence**. However, the converse is not true as the following example indicates.

EXAMPLE 11. Consider an experiment of making license plate numbers consist of 3 different letters with the sample space $\Omega = abc, bac, cab, bca, acb, cba, aaa, bbb, ccc$ where each outcome is equally likely. Define events $C_k = c$ in k^{th} position, $k = 1, 2, 3$. It is easy to see that $P(C_k) = 1/3, k = 1, 2, 3$ and $P(C_j \cap C_k) = 1/9, j \neq k$; thus C_j and C_k are independent for all $j \neq k$. However, $P(C_1 \cap C_2 \cap C_3) = 1/9 \neq 1/27$ and so C_1, C_2 and C_3 are not independent. This example shows that pairwise independence does not imply mutual independence.

1.3. Problems

- (1) Consider 3 urns. Urn A contains 2 white and 4 red balls, urn B contains 8 white and 4 red balls, and urn C contains 1 white and 3 red balls. If 1 ball is selected from each urn, what is the probability that the ball chosen from urn A was white given that exactly 2 white balls were selected?
- (2) A recent college graduate is planning to take the first three actuarial examinations in the coming summer. She will take the first actuarial exam in June. If she passes that exam, then she will take the second exam in July, and if she also passes that one, then she will take the third exam in September. If she fails an exam, then she is not allowed to take any others. The probability that she passes the first exam is .9. If she passes the first exam, then the conditional probability that she passes the second one is .8, and if she passes both the first and the second exams, then the conditional probability that she passes the third exam is .7. What is the probability that she passes all three exams?
- (3) An ectopic pregnancy is twice as likely to develop when the pregnant woman is a smoker as it is when she is a nonsmoker. If 32 percent of women of childbearing age are smokers, what percentage of women having ectopic pregnancies are smoker
- (4) A total of 46 percent of the voters in a certain city classify themselves as Independents, whereas 30 percent classify themselves as Liberals and 24 percent say that they are Conservatives. In a recent local election, 35 percent of the Independents, 62 percent of the Liberals, and 58 percent of the Conservatives voted. A voter is chosen at random? Given that this person voted in the local election, what is the probability that he or she is (a) an Independent? (b) a Liberal? (c) a Conservative? (d) What fraction of voters participated in the local election?
- (5) Consider two boxes, one containing 1 black and 1 white marble, the other 2 black and 1 white marble. A box is selected at random, and a marble is drawn from it at random. What is the probability that the marble is black? What is the probability that the first box was the one selected given that the marble is white?
- (6) If $P(A|C) > P(B|C)$ and $P(A|C^c) > P(B|C^c)$ either prove that $P(A) > P(B)$.
- (7) Let $P(A) = 0.7$ and $P(B) = 0.8$. Find $P(B)$, if a. A and B disjoint event, b. A and B independent, c. if $P(A|B) = 0.6$