6.3 확률의 성질

확률은 여러 성질을 가진다. 이 중에서 앞으로 많이 사용될 몇 가지 성질을 살펴보자. 이번에 다룰 성질들은 확률의 정의 자체가 아니라 정의로부터 유도된 것들이라는 것에 주의해야 한다. 이 장에서는 다음과 같은 확률 사건의 예를 들어 설명을 한다.

우선 사람이 20명 있는 집합이 있다고 가정하자. 이 모임에는 여자도 있고 남자도 있다. 또 성별에 관계없이 머리카락이 짧은 사람과 머리카락이 긴 두 그룹으로 나눌 수 있다고 가정한다. 여기에서 전체 20명의 집합은 표본공간이고 사람 한 명은 표본 하나다. 표본공간으로부터 0명 이상의 사람을 선택한다면 선택된 사람의 집합을 사건이라고 부르게 된다. 예를 들어 다음과 같은 사건을 생각해 보자.

- 전체 사람 중에서 남자만의 모임은 부분집합이므로 사건이라고 부를 수 있다. 이를 사건 A라고 한다.
- 전체 사람 중에서 머리카락이 긴 사람만의 모임도 부분집합이므로 사건이라고 부를 수 있다. 이를 사건 B라고 한다.
- 생일이 i월인 사람의 집합도 부분집합이므로 사건이다. 예를 들어 생일이 1월인 사람의 집합은 C_1, 생일이 2월인 사람의 집합은 C_2, ... 이런 식으로 C_{12}까지의 집합을 만들 수 있다.

원소 개수가 반드시 1보다 커야지 사건(부분집합)이 되는 것이 아니다. 원소 개수가 0인 경우, 즉 원소가 없는 부분집합도 사건이다. 예를 들어 표본공간에 생일이 6월인 사람이 존재하지 않는다면 C_6의 원소 개수는 0이다.

표본공간과 표본, 사건이 정의되었으니 각 사건에 얼마의 확률을 할당할지 정해야 한다. 여기에서는 사건에 속하는 표본 수에 비례하도록 확률을 정의해보자. 예를 들어 어떤 사건에 10명의 사람이 있다면 이 사건의 확률은 10/20 = 1/2이 된다. 이렇게 정의하면 '확률'이라는 말을 '사람 수'로 바꾸어서 생각할 수 있다. 하지만 앞에서도 강조했듯이 이 방법은 확률을 정의하는 방법 중 하나일 뿐이고 반드시 확률이란 것이 이런 식으로 정해져야 하는 것은 아니라는 점을 명심한다.

연습 문제 6.3.1

위에서 언급된 사건(A, B, C)의 표본 집합과 사건의 원소 개수와 확률을 마음대로 정해본다. 이 값은 다음 연습 문제에서 사용하게 된다.

성질 1. 공집합의 확률

> **공집합인 사건의 확률은 0이다.**

$$P(\emptyset) = 0 \tag{6.3.1}$$

(증명)

확률의 정의로부터 사건 A와 사건 B가 공통원소가 없다면 $P(A \cup B) = P(A) + P(B)$이 된다. $B = \emptyset$인 경우 A와 B의 공통원소는 없으며 $A \cup \emptyset = A$라는 사실을 이용하면

$$P(A \cup \emptyset) = P(A) = P(A) + P(\emptyset) \tag{6.3.2}$$

$$\therefore P(\emptyset) = 0 \tag{6.3.3}$$

성질 2. 여집합의 확률

어떤 사건의 여집합인 사건의 확률은 (1 - 원래 사건의 확률)과 같다.

$$P(A^C) = 1 - P(A) \tag{6.3.4}$$

어떤 사건의 여집합이란 그 사건의 표본이 아닌 표본의 집합을 말하며 사건 기호에 C라는 윗첨자를 붙여서 표시한다. 위의 예에서는 A^C는 남자라는 부분집합 A에 대한 여집합이므로 여자 집합이 된다.

(증명)

확률의 정의로부터 사건 A와 사건 B가 공통원소가 없다면 $P(A \cup B) = P(A) + P(B)$이 된다.
$B = A^C$인 경우 A와 B의 공통원소는 없다.

$$P(A \cup A^C) = P(\Omega) = 1 = P(A) + P(A^C) \tag{6.3.5}$$

$$\therefore P(A^C) = 1 - P(A) \tag{6.3.6}$$

위 성질과 $P(A^C)$가 0보다 크거나 같아야 한다는 콜모고로프의 공리 1을 결합하면 **확률값은 0과 1 사잇값을 가져야 한다.**

$$P(A^C) = 1 - P(A) \geq 0 \tag{6.3.7}$$

$$0 \leq P(A) \leq 1 \tag{6.3.8}$$

성질 3. 포함-배제 원리

두 사건의 합집합의 확률은 각 사건의 확률의 합에서 두 사건의 교집합의 확률을 뺀 것과 같다.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{6.3.9}$$

이를 **포함-배제 원리(Inclusion-exclusion principle)** 혹은 덧셈 규칙(sum rule, addition law)이라 한다.

위의 예에서 사건 A와 사건 B의 교집합이란 남자의 집합에 속하면서 머리카락이 긴 사람의 집합에 속하는 사람의 집합 즉, 장발 남자 집합을 말한다.

(증명)

$$\begin{aligned} P(A \cup B) &= P(A \cup (B \cap A^C)) \\ &= P(A) + P(B \cap A^C) \\ &= P(A) + P(B \cap A^C) + P(A \cap B) - P(A \cap B) \\ &= P(A) + P((A^C \cap B) \cup (A \cap B)) - P(A \cap B) \\ &= P(A) + P(B) - P(A \cap B) \end{aligned} \tag{6.3.10}$$

성질 4. 전체 확률의 법칙

복수의 사건 C_i가 다음을 만족하는 사건들이라고 가정한다.

- 서로 교집합이 없다. 이를 서로 배타적(mutually exclusive)이라고도 한다.

$$C_i \cap C_j = \emptyset \quad (i \neq j) \tag{6.3.11}$$

- 모든 집합의 합집합이 전체집합(표본공간)이다. 이 경우 완전한(complete) 부분집합들이라고 한다.

$$C_1 \cup C_2 \cup \cdots = \Omega \tag{6.3.12}$$

이 경우,

> **사건 A의 확률은 사건 A와 사건 C_i가 동시에 발생할 사건들의 확률의 합과 같다.**

$$P(A) = \sum_i P(A \cap C_i) \tag{6.3.13}$$

이를 **전체 확률의 법칙(law of total probability)**이라 한다.

위의 예를 사용하면 사건 $P(A \cap C_1)$은 생일이 1월인 남자의 집합이고 전체 확률의 법칙은 다음과 같이 이해할 수 있다.

> 생일이 1월인 남자라는 사건의 확률과 생일이 2월인 남자라는 사건의 확률, 이 이외에도 각각의 월을 생일로 가지는 남자라는 사건들의 확률을 모두 합치면 남자라는 사건의 확률이 된다.

(증명)

$$\begin{aligned} A &= A \cap \Omega \\ &= A \cap (C_1 \cup C_2 \cup \cdots) \\ &= (A \cap C_1) \cup (A \cap C_2) \cup \cdots \end{aligned} \tag{6.3.14}$$

C_i가 서로 공통 원소가 없기 때문에 $A \cap C_i$도 서로 공통 원소가 없다. 따라서 확률의 정의에 따라 다음 등식이 성립한다.

$$P(A) = P(A \cap C_1) + P(A \cap C_2) + \cdots = \sum_i P(A \cap C_i) \tag{6.3.15}$$

교집합의 확률은 다음처럼 쉼표로 표시하기도 한다.

$$P(A \cap B) = P(A, B) \tag{6.3.16}$$

따라서 전체확률의 법칙은 다음처럼 쓸 수 있다.

$$P(A) = \sum_i P(A, C_i) \tag{6.3.17}$$

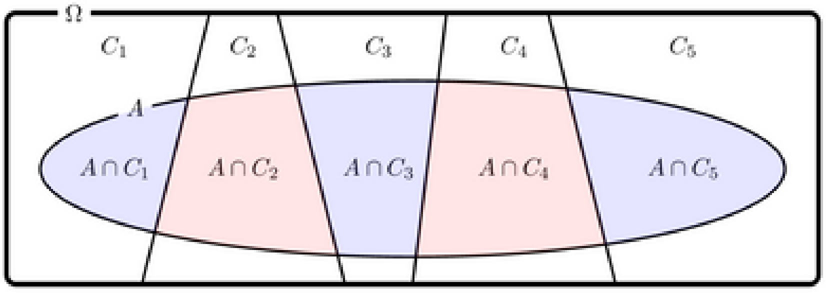

그림 6.3.1 : 전체확률의 법칙

연습 문제 6.3.2

앞 연습 문제에서 정한 확률 값을 이용하여 확률의 성질 2, 3, 4번이 성립하는지 확인해본다.

지금까지 공부한 확률의 성질을 정리하면 다음과 같다.

확률의 성질 요약

- 공집합의 확률

$$P(\emptyset) = 0 \tag{6.3.18}$$

- 여집합의 확률

$$P(A^C) = 1 - P(A) \tag{6.3.19}$$
$$0 \leq P(A) \leq 1 \tag{6.3.20}$$

- 포함-배제 원리

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{6.3.21}$$

- 전체 확률의 법칙

$$C_i \cap C_j = \emptyset \quad (i \neq j) \tag{6.3.22}$$
$$C_1 \cup C_2 \cup \cdots = \Omega \tag{6.3.23}$$
$$P(A) = \sum_i P(A, C_i) \tag{6.3.24}$$