

6.2 확률의 수학적 정의와 의미

일반적으로 확률을 다룰 때는 현실에서 해결하고자 하는 문제와 결부하여 정의한다. 예를 들어 다음과 같은 문제들을 생각해보자.

1. 동전을 한번 던졌다. 동전이 앞면이 나올 것인가 뒷면이 나올 것인가?
2. 플레잉카드(트럼프카드) 뭉치에서 카드 한장을 선택했다. 이 카드는 어떤 무늬인가?
3. 동전을 두 번 던져 나올 수 있는 경우는?
4. 약속 날짜를 정하기로 했다. 결정된 날짜가 31일인가 아닌가?
5. 과일가게에서 손님이 과일을 하나 샀다. 이 과일은 어떤 과일일까?
6. 삼성전자 주식의 가격은 내일 몇 % 오를까?
7. 회전하는 원판에 화살을 쏘았다. 화살이 박힌 위치의 각도는 기준선에서 몇 도 위치인가?
8. 체온을 측정했다. 체온이 몇 도일까?

이 문제들의 공통점은 답을 100% 확신할 수 없다는 점이다. 어떤 문제는 무엇이 답인지 전혀 예측할 수 없는 것도 있고 어떤 문제는 어느 정도의 정확도 혹은 범위 내에 있다고 예측할 수 있는 것도 있다. 확률론은 이러한 문제가 어떤 답을 가질 수 있고 그 답의 신뢰성이 얼마나 되는지 계산하는 정량적인 방법을 제시한다.

이 장에서는 위와 같은 문제를 풀기 위해 확률의 수학적 정의를 살펴보는 것부터 시작한다. 확률의 수학적 정의를 공부하는 것은 확률과 관련된 복잡한 문제를 푸는 것뿐 아니라 지금까지 생각해오던 확률에 대한 관점을 새롭게 환기시키는데도 큰 도움이 된다.

확률을 수학적으로 정의하려면 다음과 같은 3가지 개념을 알아야 한다.

- 확률표본
- 표본공간
- 사건

표본공간과 확률표본

우선 확률표본(sample)과 표본공간(sample space)을 정의한다. **확률표본(probabilistic sample, random sample)** 또는 **표본(sample)**은 풀고자 하는 확률적 문제에서 **발생(realize)할 수 있는 하나의 현상**, 혹은 **선택(sampled)될 수 있는 하나의 경우**를 말한다. **표본공간(sample space)**은 가능한 모든 표본의 집합을 말한다. 보통 Ω (대문자 오메가)라는 그리스 문자로 표기한다. 표본공간을 정의한다는 것은 우리가 고려하는 범위에서 **어떤 표본(경우, 현상)이 가능하고 어떤 표본이 가능하지 않은가**를 정의하는 작업이다.



그림 6.2.1 : 표본과 표본공간

예제

예를 들어 동전 던지기 문제에서는 "앞면(Head)이 나오는 현상" 또는 "뒷면(Tail)이 나오는 결과"가 각각 표본이 될 수 있다. 이 표본들을 각각 기호 "H"와 "T"로 표시하기로 하자. 동전이 세로로 서 있는 경우나 동전이 대각선으로 서 있는 경우 등을 고려할지 고려하지 않을지는 동전 던지기라는 물리현상에 대한 지식과 문제의 필요성에 의해 달라진다. 우리는 일상에서 경험한 지식(데이터)으로 동전이 세로로 서 있는 경우나 동전이 대각선으로 서 있

는 경우가 극히 드물다는 것을 알고 있기 때문에 이러한 경우는 고려할 필요가 없다고 가정할 수 있다. 이때는 표본공간이라는 집합이 "H"와 "T" 두 가지 원소(표본)로만 구성된다. 따라서 1번 예제의 표본공간은 다음처럼 표기할 수 있다.

$$\Omega_1 = \{H, T\} \quad (6.2.1)$$

예제

플레이카드 문제에서는 스페이드(♠), 하트(♥), 다이아몬드(♦), 클럽(♣) 기호를 이용하여 표본을 표시할 수 있다. 따라서 2번 예제의 표본공간은 다음처럼 표기할 수 있다.

$$\Omega_2 = \{\spadesuit, \heartsuit, \diamondsuit, \clubsuit\} \quad (6.2.2)$$

연습 문제 6.2.1

- (1) 위에서 예로 든 문제 중 동전을 두 번 던지는 문제를 확률론적으로 접근할 때 표본공간 Ω_3 을 구하라.
- (2) 위에서 예로 든 문제 중 약속 날짜 문제에서 결정된 날짜가 31일인가 아닌지 구하는 문제를 확률론적으로 접근할 때 표본공간 Ω_4 을 구하라.

예제

표본공간은 풀고자 하는 문제에 대한 우리의 지식이나 필요성에 따라 달라진다. 예를 들어 과일가게에서 손님이 산 과일을 맞추는 예제에서 표본공간은 그 과일가게에서 어떤 과일을 파는가에 따라 달라진다. 예를 들어 과일가게 문제에서 사과("A": Apple)와 오렌지("O": Orange)만 파는 과일가게라면 표본공간은 다음과 같다.

$$\Omega_5 = \{A, O\} \quad (6.2.3)$$

예제

표본이 연속적인 숫자이면 표본공간이 무한대의 원소를 가질 수 있다. 예를 들어 삼성전자 주식의 가격 문제에서 주가는 한국거래소의 규정에 따라 다음날 -30%부터 30%까지 내리거나 오를 수 있다. 따라서 표본공간은 다음과 같다.

$$\Omega_6 = \{x : -30 \leq x \leq 30\} \quad (6.2.4)$$

이 표본공간은 구간 내의 모든 실수를 원소로 가지므로 원소의 개수가 무한대다.

(주1: 한국거래소의 규정에 따라 틱사이즈(tick size)라고 부르는 가격 단위가 정해져 있으므로 엄격히 말하면 원소 개수가 무한대는 아니다. 그러나 여기에서는 이를 무시하자.)

연습 문제 6.2.2

회전하는 원판의 각도를 결정하는 문제의 표본공간 Ω_7 는? 이 표본공간의 표본 개수는?

예제

체온을 결정하는 문제를 살펴보자. 만약 정상적인 사람이라면 체온은 보통 36도에서 40도 사이일 것이다. 그러나 표본공간은 가능성이 0이 아닌 모든 경우를 포함해야 하므로 36도보다 낮은 경우, 혹은 40도보다 높은 경우가 **절대로 없다고** 말할 수는 없다. 따라서 이런 경우에는 수학적 편의를 위해 실수 전체의 집합을 표본공간으로 잡는 것이 편리하다.

$$\Omega_8 = \mathbf{R} \quad (6.2.5)$$

연습 문제 6.2.3

확률론으로 접근할 수 있는 문제와 그 표본공간의 예를 4가지 더 만들어본다. 4가지 문제 중 2가지는 표본공간의 크기가 무한대이어야 한다.

사건

사건(event)는 표본공간 Ω 의 부분집합, 즉, 전체 표본공간 중에서 우리가 관심을 가지고 있는 일부 표본의 집합을 뜻한다. 보통 A, B, C, \dots 식으로 대문자 알파벳으로 표기한다.

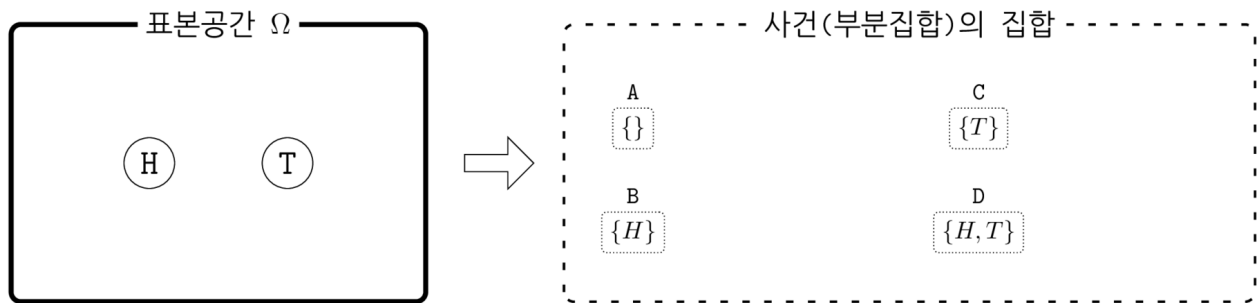


그림 6.2.2 : 사건

위에서 예로 든 동전 표본공간에서는 가능한 사건(부분집합)은 다음과 같다.

$$A = \{\} = \emptyset \quad (6.2.6)$$

$$B = \{H\} \quad (6.2.7)$$

$$C = \{T\} \quad (6.2.8)$$

$$D = \{H, T\} = \Omega \quad (6.2.9)$$

예로 든 B 라는 사건은 ‘동전의 앞면이 나오는 경우’를 뜻하고 D 라는 사건은 ‘동전의 앞면이 나오거나 뒷면이 나오는 경우’를 뜻한다. ‘~가 나오는 경우’라고 할 때 이 ‘경우’라는 개념이 바로 사건(부분집합)에 해당한다.

파이썬으로 다음처럼 부분집합의 집합을 구현할 수 있다. 부분집합을 set 자료형이 아닌 frozenset 자료형으로 만든 이유는 딕셔너리의 키(key)로 사용하기 위해서다.

In [1]:

```
A = frozenset([])
B = frozenset(['H'])
C = frozenset(['T'])
D = frozenset(['H', 'T'])
set([A, B, C, D])
```

Out[1]:

```
{frozenset(), frozenset({'H'}), frozenset({'T'}), frozenset({'H', 'T'})}
```

연습 문제 6.2.4

- (1) 플레이카드 한장을 뽑아서 무늬를 결정하는 문제의 표본공간 Ω_2 의 모든 사건을 구하고 이를 frozenset 의 set 으로 만든다.
- (2) 동전을 두 번 던지는 문제의 표본공간 Ω_3 의 모든 사건을 구하고 이를 frozenset 의 set 으로 만든다.

확률

확률(probability)이란 **사건(부분집합)**을 입력하면 **숫자(확률값)**가 출력되는 함수다.

$$\text{사건(부분집합)} \xrightarrow{\text{확률}} \text{숫자(확률값)} \quad (6.2.10)$$

확률이라는 함수의 입력값의 집합 즉, 정의역(domain)은 표본공간의 모든 사건(부분집합)의 집합이다.

(주2: 엄격한 정의로는 모든 사건의 집합이 아니라 시그마대수(sigma algebra)라는 특별한 사건 집합에 대해서만 정의하면 된다. 하지만 여기에서는 시그마 대수는 고려하지 않기로 한다.)

즉, **모든 각각의 사건(부분집합)에 어떤 숫자를 할당(assign, allocate)하는 함수가 확률**이다. 보통 대문자 알파벳 P 로 나타낸다. 확률 P 는 함수이고 $P(A)$ 는 A 라는 사건(부분집합)에 할당된 숫자를 뜻한다.

다만 이 함수는 다음과 같은 세가지 규칙을 지켜야 한다.

- (1) 모든 사건에 대해 확률은 실수이고 0 또는 양수다.

$$P(A) \geq 0 \quad (6.2.11)$$

- (2) 표본공간(전체집합)이라는 사건(부분집합)에 대한 확률은 1이다.

$$P(\Omega) = 1 \quad (6.2.12)$$

- (3) 공통 원소가 없는 두 사건의 합집합의 확률은 사건별 확률의 합이다.

$$A \cap B = \emptyset \rightarrow P(A \cup B) = P(A) + P(B) \quad (6.2.13)$$

이 세 가지를 **콜모고로프의 공리(Kolmogorov's axioms)**라고 한다.

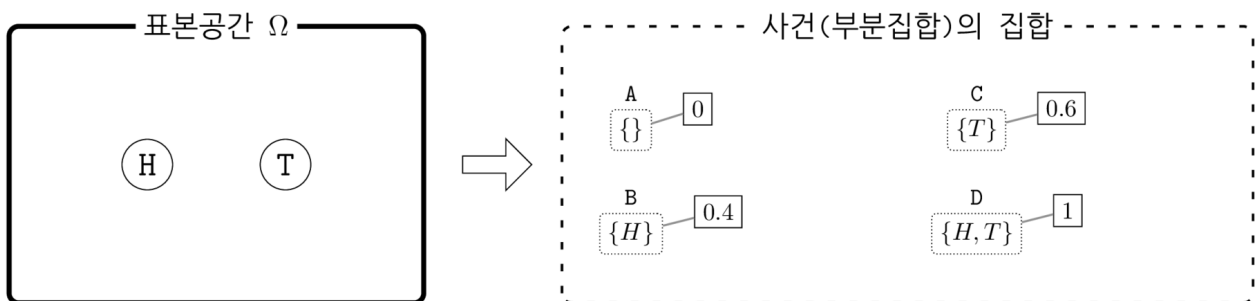


그림 6.2.3 : 확률

예를 들어 표본공간이 $\Omega = \{H, T\}$ 일 때 사건(부분집합) $\{H\}$ 에 대한 확률값 $P(\{H\})$ 는 "H라는 표본이 선택될 확률"이고 사건(부분집합) $\{H, T\}$ 에 대한 확률값 $P(\{H, T\})$ 는 "H 또는 T라는 표본이 선택될 확률"이다. 따

라서 확률이라는 함수를 정의한다는 것은 "무엇이 선택될 확률이 얼마인가"라는 질문에 대한 답을 모든 경우(사건, 부분집합)에 대해 미리 준비해놓은 것 또는 할당해놓은 것이라 할 수 있다.

확률은 표본이 아닌 사건을 입력으로 가지는 함수

확률에 대해 흔히 가지고 있는 잘못된 지식은 확률이 "표본 하나 하나에 대해 정의되어 있는 숫자"라는 것이다. 즉, 확률이 표본을 입력받아 숫자(확률값)를 출력하는 함수라고 잘못 알고 있다. 예를 들어 주사위를 던져 숫자 1이 나타나는 경우에 대해 다음처럼 생각한다.

$$P(1) = \frac{1}{6} \quad (6.2.14)$$

하지만 위 식은 틀린 것이다. 확률은 표본이 아닌 **사건(부분집합)에 대해** 정의하기 때문이다. 올바른 식은 다음과 같다.

$$P(\{1\}) = \frac{1}{6} \quad (6.2.15)$$

파이썬을 이용한 확률의 구현

파이썬에서는 사건을 키(key)로 가지고 숫자를 값(value)으로 가지는 딕셔너리를 사용하여 확률 함수를 구현할 수 있다. 이때 딕셔너리의 키는 가능한 모든 부분집합이 되어야 한다. 동전의 경우에는 가능한 부분집합 수가 $A = \emptyset, B = \{H\}, C = \{T\}, D = \{H, T\}$ 4개이므로 파이썬으로 구현하면 다음과 같다.

In [2]:

```
P = {A: 0, B: 0.4, C: 0.6, D: 1}
P
```

Out[2]:

```
{frozenset(): 0,
 frozenset({'H'}): 0.4,
 frozenset({'T'}): 0.6,
 frozenset({'H', 'T'}): 1}
```

이 예에서 동전의 앞면이 나오는 경우에 대해 할당된 확률값이 0.5가 아닌 0.4임에 주의하라. 콜모고로프의 공리만 지킨다면 각각의 사건(부분집합)에 대한 확률값은 어떤 값도 할당해도 된다.

위와 같은 할당법이 콜모코로프의 공리를 지키고 있는지 확인해보자.

(1) 모든 확률은 0 이상이다.

$$P(\emptyset) = 0 \geq 0 \quad (6.2.16)$$

$$P(\{H\}) = 0.4 \geq 0 \quad (6.2.17)$$

$$P(\{T\}) = 0.6 \geq 0 \quad (6.2.18)$$

$$P(\{H, T\}) = 1 \geq 0 \quad (6.2.19)$$

따라서 첫 번째 공리를 만족한다.

(2) 전체집합에 대한 확률은 1이다.

$$P(\Omega) = P(\{H, T\}) = 1 \quad (6.2.20)$$

따라서 두 번째 공리를 만족한다.

(3) 교집합이 공집합인 사건(부분집합)의 합집합인 사건의 확률은 각 사건(부분집합)의 확률의 합이다.

$$P(\{H\}) = 0.4 = P(\{H\} \cup \emptyset) = 0.4 + 0 \quad (6.2.21)$$

$$P(\{T\}) = 0.6 = P(\{T\} \cup \emptyset) = 0.6 + 0 \quad (6.2.22)$$

$$P(\{H, T\}) = 1 = P(\{H, T\} \cup \emptyset) = 1 + 0 \quad (6.2.23)$$

$$P(\{H, T\}) = 1 = P(\{H\} \cup \{T\}) = 0.4 + 0.6 \quad (6.2.24)$$

교집합이 공집합인 어떤 사건을 선택하더라도 합집합의 확률은 각 사건의 확률의 합과 같다. 따라서 세 번째 공리를 만족한다.

따라서 위에서 우리가 정의한 함수는 세 공리를 모두 만족시키는 할당법이므로 확률이라고 부를 수 있다.

확률은 골동품을 포장해서 파는 가게의 가격에 비유할 수 있다.

- 가게에서 파는 하나 하나의 골동품은 표본이다. 모든 골동품은 서로 다르기 때문에 똑같은 물건(표본)은 없다.
- 가게에서 파는 모든 골동품의 집합은 표본공간(전체집합)이다.
- 사건(부분집합)은 골동품을 넣은 상자를 말한다. 상자안의 골동품 개수에는 제한이 없다. 상자안의 골동품이 하나가 될 수도 있고 골동품이 없는 빈 포장(공집합)도 가능하다.
- 확률은 상자에 붙인 가격표 숫자를 말한다. 가격은 마음대로 붙여도 되지만 다음 규칙을 지켜야 한다.
 1. 음수인 가격은 없다. 공짜(0)나 양수이어야 한다.
 2. 가게안의 모든 골동품을 하나의 상자에 포장하면 그 상자의 가격은 1이다.
 3. 공통적으로 포함된 골동품이 없는 두개의 상자의 가격은 그 두개의 포장에 들어간 골동품을 합쳐서 하나의 상자로 만들었을 때의 가격과 같아야 한다. 즉 상자를 나누거나 합쳤다고 가격이 달라져서는 안 된다.

연습 문제 6.2.5

(1) 플레이카드 한 장을 뽑아서 무늬를 결정하는 문제에 대해 위와 같이 파이썬으로 확률을 할당해본다.

(2) 동전을 두 번 던지는 문제에 대해 위와 같이 파이썬으로 확률을 할당해본다.

주사위 한 면의 확률은 무조건 $\frac{1}{6}$?

우리는 주사위의 어떤 한 면이 나올 확률을 $\frac{1}{6}$ 이라고 배웠다. 하지만 이 값은 주사위 확률의 하나의 예(example)에 지나지 않으며 다른 확률도 가능하다.

여러분 손에 현실의 주사위가 있다고 하자. 그 주사위를 던지면 정말로 특정한 하나의 면이 나올 가능성이 "정확하게" 6분의 1일까? 이렇게 묻는다면 대부분은 "정확하게는 그렇지 않다"고 대답할 것이다. 그 이유는 주사위를 아주 정밀하게 만들지 않으면 어떤 면이 미세하게나마 더 잘 나오게 되는 경우가 발생할 수 있다는 사실을 알고 있기 때문이다. 더 극단적인 경우를 생각해보자. 누군가가 주사위 도박에서 사기를 치려고 특수한 주사위를 만들었다고 하자. 이 주사위는 절대로 6이 나오지 않으며 10번을 던지면 5번은 반드시 1이 나오게 만들었다면? 이 때는 확률값이 어떻게 할당되어 있는 상태인가? 이때는 확률이 다음과 같이 할당되어 있다고 할 수 있다.

$$P(\{1\}) = 0.5 \quad (6.2.25)$$

$$P(\{6\}) = 0 \quad (6.2.26)$$

이 두 사건을 포함한 다른 모든 사건의 확률이 콜모고로프 정의를 만족한다면 이러한 확률값도 확률의 수학적 정의에 맞다고 할 수 있다.

간단한 예를 들어 1부터 4까지의 숫자만 나오는 4면체 주사위를 생각하자. 이 주사위가 공평한 주사위라면 확률은 다음처럼 할당된다.

$$\begin{aligned} P(\emptyset) &= 0, \\ P(\{1\}) &= \frac{1}{4}, \quad P(\{2\}) = \frac{1}{4}, \quad P(\{3\}) = \frac{1}{4}, \quad P(\{4\}) = \frac{1}{4}, \\ P(\{1, 2\}) &= \frac{1}{2}, \quad P(\{1, 3\}) = \frac{1}{2}, \quad P(\{1, 4\}) = \frac{1}{2}, \\ P(\{2, 3\}) &= \frac{1}{2}, \quad P(\{2, 4\}) = \frac{1}{2}, \quad P(\{3, 4\}) = \frac{1}{2}, \\ P(\{1, 2, 3\}) &= \frac{3}{4}, \quad P(\{1, 2, 4\}) = \frac{3}{4}, \quad P(\{1, 3, 4\}) = \frac{3}{4}, \quad P(\{2, 3, 4\}) = \frac{3}{4}, \\ P(\{1, 2, 3, 4\}) &= 1 \end{aligned}$$

하지만 절대로 1이 나오지 않도록 조작된 4면체 주사위라면 다음과 같이 확률이 할당되었을 수도 있다.

$$\begin{aligned} P(\emptyset) &= 0 \\ P(\{1\}) &= 0, \quad P(\{2\}) = \frac{1}{2}, \quad P(\{3\}) = \frac{1}{4}, \quad P(\{4\}) = \frac{1}{4}, \\ P(\{1, 2\}) &= \frac{1}{2}, \quad P(\{1, 3\}) = \frac{1}{4}, \quad P(\{1, 4\}) = \frac{1}{4}, \\ P(\{2, 3\}) &= \frac{3}{4}, \quad P(\{2, 4\}) = \frac{3}{4}, \quad P(\{3, 4\}) = \frac{1}{2}, \\ P(\{1, 2, 3\}) &= \frac{3}{4}, \quad P(\{1, 2, 4\}) = \frac{3}{4}, \quad P(\{1, 3, 4\}) = \frac{3}{4}, \quad P(\{2, 3, 4\}) = 1, \\ P(\{1, 2, 3, 4\}) &= 1 \end{aligned} \quad (6.2.28)$$

이 확률 할당은 콜모코로프의 공리를 모두 만족하고 있다.

1. 각 확률은 모두 0보다 같거나 크다.
2. 전체집합에 할당된 확률은 1이다.
3. 교집합이 공집합인 모든 두 부분집합의 확률은 각 부분집합의 확률의 합이다.

$$P(\{1, 2\}) = \frac{1}{2} = P(\{1\}) + P(\{2\}) = 0 + \frac{1}{2}$$

$$P(\{1, 3\}) = \frac{1}{4} = P(\{1\}) + P(\{3\}) = 0 + \frac{1}{4}$$

$$P(\{1, 4\}) = \frac{1}{4} = P(\{1\}) + P(\{4\}) = 0 + \frac{1}{4}$$

$$P(\{2, 3\}) = \frac{3}{4} = P(\{2\}) + P(\{3\}) = \frac{1}{2} + \frac{1}{4}$$

$$P(\{2, 4\}) = \frac{3}{4} = P(\{2\}) + P(\{4\}) = \frac{1}{2} + \frac{1}{4}$$

$$P(\{3, 4\}) = \frac{1}{2} = P(\{3\}) + P(\{4\}) = \frac{1}{4} + \frac{1}{4}$$

$$P(\{1, 2, 3\}) = \frac{3}{4} = P(\{1, 2\}) + P(\{3\}) = \frac{1}{2} + \frac{1}{4}$$

$$P(\{1, 2, 3\}) = \frac{3}{4} = P(\{1\}) + P(\{2, 3\}) = 0 + \frac{3}{4}$$

⋮

(6.2.29)

연습 문제 6.2.6

두 조건

$$P(\{1\}) = 0.5 \quad (6.2.30)$$

$$P(\{6\}) = 0 \quad (6.2.31)$$

을 만족하도록 주사위의 확률을 모든 사건(부분집합)에 대해 할당하고 이를 파이썬으로 구현한다.

그렇다면 우리는 왜 지금까지 주사위의 한 면이 나올 확률은 무조건 $\frac{1}{6}$ 이라고 생각해 왔던 것일까? 그 이유는 확률의 정의와는 관계없는 다음 사실을 가정하고 있기 때문이다.

주사위가 **공정(fair)한 주사위**이다. 혹은 공정하지 않다고 생각할 수 있는 증거가 아직 없다.

주사위가 공정하다면 특정한 주사위 하나의 면이 나올 확률값은 모두 같고 이를 x 라 하자.

$$P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\}) = P(\{5\}) = P(\{6\}) = x \quad (6.2.32)$$

그리고 이 사건들은 서로 공통원소가 없고 합집합은 전체 집합이므로 확률의 성질을 이용하여 다음과 같이 각 확률이 $\frac{1}{6}$ 임을 구할 수 있다.

$$\begin{aligned} P(\{1\}) + P(\{2\}) + P(\{3\}) + P(\{4\}) + P(\{5\}) + P(\{6\}) &= 6x \\ &= P(\{1, 2, 3, 4, 5, 6\}) \\ &= P(\Omega) = 1 \end{aligned} \quad (6.2.33)$$

$$\therefore x = 1/6 \quad (6.2.34)$$

즉, 이 방법은 확률 값을 만드는 방법의 하나일 뿐이고 현실에서 꼭 이대로 확률 값 즉, 숫자를 배당할 이유는 없다.

이 논리에 따르면 표본의 개수가 유한하고 각 사건에 대해 원소의 개수 이외의 아무런 정보가 없다면 각 사건의 확률을 다음과 같다고 보는 것이 타당하다.

$$P(A) = \frac{\text{card}(A)}{\text{card}(\Omega)} \quad (6.2.35)$$

그러나 만약 다른 곳에서 얻은 정보(예 : 자연법칙 또는 도메인 지식)이나 표본에 대한 데이터가 존재하면 더 믿을 수 있는 확률값을 가정할 수도 있다.

연습 문제 6.2.7

- (1) 약속 날짜가 31일인가 아닌가를 결정하는 문제에서 확률을 할당해보자.
- (2) 사과와 오렌지만 파는 과일가게에서 손님이 선택한 과일이 어떤 과일인지 결정하는 문제에 대해 확률을 할당해보자.
- (3) 사과와 오렌지만 파는 과일가게에 과일이 100개가 진열되어 있고 이 중 70개가 사과, 30개가 오렌지이다. 손님이 선택한 과일이 어떤 과일인지 결정하는 문제에 대해 확률을 할당해보자.

확률의 의미

지금까지 우리는 표본 집합의 부분 집합인 사건에 대해 확률값이라는 숫자를 할당했다. 이 확률값이라는 숫자는 도대체 어떤 의미를 가지는 걸까? 확률값이라는 숫자가 가지는 의미에 대해서는 여러 해석이 있을 수 있다. 그중 가장 대표적인 것이 **빈도주의(Frequentist) 관점**과 **베이지안(Bayesian) 관점**이다.

빈도주의에서는 **반복적으로 선택된 표본이 사건(부분 집합) A의 원소가 될 경향(propensity)**을 그 사건의 확률이라고 본다. 예를 들어 동전을 던져 "앞면이 나오는 사건"의 확률값이 0.5라는 것은 빈도주의 관점에서는 실제로 동전을 반복하여 던졌을 경우 동전을 던진 전체 횟수에 확률값을 곱한 숫자만큼 해당 사건이 발생한다고 본다. 예를 들어 10,000번을 던지면 $10,000 \times 0.5 = 5,000$ 번 앞면이 나오는 경향을 가진다는 의미다.

베이지안 관점에서 확률은 **"선택된 표본이 특정한 사건(부분 집합)에 속한다는 가설(hypothesis), 명제(proposition) 혹은 주장(assertion)의 신뢰도(degree of belief)"**라고 본다. 반복이라는 개념은 사용되지 않는다.

예를 들어

새는 날 수 있다

라는 명제가 있다고 하자. 이 명제는 참인가? 대부분의 새는 날 수 있으므로 이 명제는 참에 가깝다. 하지만 닭, 타조, 펭귄 등의 새나 혹은 어린 새, 날개를 다친 새는 날 수 없으므로 항상 참인 명제는 아니다. 그렇다고 명제 전체를 거짓이라고 하는 것은 실용적이지 못하다. 이런 경우 숫자를 사용하여 "이 명제의 신뢰도는 95%이다"라고 명시할 수 있다면 유용할 것이다. 베이지안 관점에서의 확률은 이러한 명제의 신뢰도를 뜻한다. 따라서 베이지안 관점에서는 위 명제를 다음처럼 서술한다.

새가 날 수 있는 가능성은 95%이다

위의 동전 문제에서 동전을 던져

"앞면이 나오는 사건"의 확률값이 0.5다

라는 말의 의미는 베이지안 관점에서는

"앞면이 나왔다"는 주장의 신뢰도가 0.5이다

라는 의미다.

4지 선다형 객관식 문제를 풀 때도 우리는 베이지안 확률을 사용한다. 1번부터 4번까지의 보기를 읽어보고 마음속으로 다음처럼 각각의 보기에 대한 확률을 할당하기 때문이다.

'1번은 절대로 답이 될 수 없어. 그러니까 1번이 정답일 확률은 0이다. 2번과 3번이 그럴 듯한데. 4번은 가능성이 2번이나 3번의 가능성의 반도 안 되어 보이고... 그러니까 2번과 3번이 정답일 확률은 각각 0.4이고 4번이 정답일 확률은 0.2이군.'

이러한 생각을 했다면 $\{1, 2, 3, 4\}$ 라는 표본 집합이 있을 때 다음과 같이 확률을 할당한 것이다.

$$\begin{aligned}P(\{1\}) &= 0 \\P(\{2\}) &= 0.4 \\P(\{3\}) &= 0.4 \\P(\{4\}) &= 0.2\end{aligned}\tag{6.2.36}$$

여기에서 확률의 의미는 무언가 반복되는 것, 또는 빈도와는 전혀 관계가 없다. 확률 $P(\{1\})$ 은 "정답이 1이다"라는 주장에 대한 신뢰도일 뿐이다.

베이지안 관점에서 사건(부분집합)이란 "원하는 답(표본)이 포함되어 있을 가능성이 있는 후보의 집합"이며 이런 맥락에서 우리가 어떤 사건을 제시하면 그 자체로 "이 사건에 속한 원소 중에 원하는 답(표본)이 있다"는 명제 혹은 주장을 제시한 것이라 할 수 있다.

또한 베이지안 확률론에서 **사건이 일어났다(occur)** 혹은 **발생했다**하는 말은 그 **사건(부분집합)의 원소 중에 정말로 선택된 표본이 있다는 사실을 알게 되었다**는 것을 말한다. 다른 말로는 해당 사건이 말하고 있는 주장이 진실임을 알게 되었다는 뜻으로 지금까지 모르고 있던 **추가적인 정보**가 들어왔음을 뜻한다.

예를 들어 불투명 컵 안에 주사위를 넣고 굴렸다고 가정하자. '주사위의 눈금이 짝수가 나오는 사건이 발생했다'라는 말은 컵을 들어서 주사위의 눈금을 보고 '주사위의 눈금이 짝수다'라는 사실을 알게 된 것을 의미한다. 이 용어는 나중에 베이지 법칙을 설명할 때 사용된다.

확률의 빈도주의적 관점과 베이지안 관점은 양립할 수 없는 관계가 아니다. 예를 들어 의사가 환자를 진찰한 후

"검진 결과로 보아 암에 걸렸을 확률이 90%이다"

라고 진단했다고 하자.

의사의 관점에서는 '이러한 검진 결과를 가진 환자를 정밀 검사로 확인하는 일을 100번 반복하면 그중의 약 90명은 암에 걸려 있다.'라는 빈도주의적 관점일 수 있다. 하지만 환자의 관점에서는 다른 환자가 암인가 아닌가는 의미가 없기 때문에 '자신이 암에 걸렸다는 의사의 주장이 사실일 가능성은 90%이다'라는 베이지안 관점으로 바라보게 된다.

