통계학

– 제 3 장 확률과 확률분포 (1)

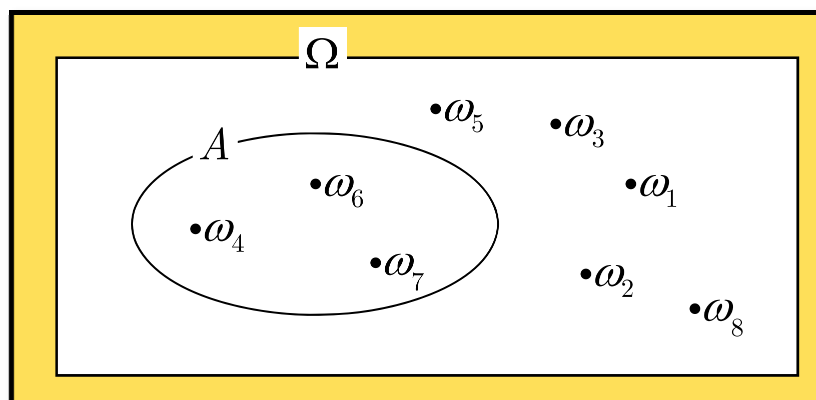
0011 0010 1010 1101 0001 0100 1011

확률의 정의

- 확률의 정의
  - 확률을 정의하기 위한 기본 개념
    - 통계적 실험(statistical experiment)
      - 출현 가능한 모든 결과들 중 오직 한 가지 결과만 나타나는 실험
      - 반복가능
    - 표본공간(sample space, Ω)
      - 통계적 실험에서 출현 가능한 모든 결과들의 집합
    - 근원사상(elementary outcome, $\omega_1,\ \omega_2,\ \cdots$)
      - 표본공간을 구성하는 개개의 결과
    - 사상(events, $A,\ B,\ \cdots$)
      - 표본공간의 부분집합
      - 어떤 특성을 갖는 결과들의 모임

- 예 – [동전 던지기]
  - 통계적 실험 : 1개의 공정한 동전을 세 번 던졌을 때 나오는 면을 기록
  - 표본공간 : $\Omega = \{\mathrm{HHH},\ \mathrm{HHT},\ \mathrm{HTH},\ \mathrm{HTT},\ \mathrm{THH},\ \mathrm{THT},\ \mathrm{TTH},\ \mathrm{TTT}\}$
  - 근원사상 : $\omega_1 = \mathrm{HHH},\ \omega_2 = \mathrm{HHT},\ \omega_3 = \mathrm{HTH},\ \omega_4 = \mathrm{HTT},$
    $\omega_5 = \mathrm{THH},\ \omega_6 = \mathrm{THT},\ \omega_7 = \mathrm{TTH},\ \omega_8 = \mathrm{TTT}$
  - 정확히 앞면이 한 번 나오는 사상

$$A = \{\ \omega_4,\ \omega_6,\ \omega_7\ \}$$
$$= \{\mathrm{HTT},\ \mathrm{THT},\ \mathrm{TTH}\}$$



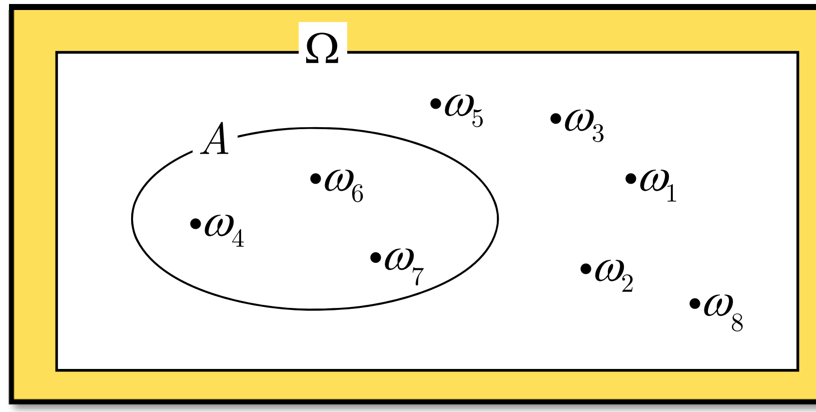
- 확률의 정의
  - 확률 (probability)
    - 각각의 근원 사상들이 발생할 가능성 같다고 가정
    - 사상 A의 확률 $P(A)$는

$$P(A) = \frac{\text{사상}\,A\text{ 에 속하는 근원사상의 수}}{\text{표본공간의 모든 근원사상의 수}}$$

  - 확률의 성질
    - 모든 사상 $A \subset \Omega$에 대하여 $0 \le P(A) \le 1$
    - $P(A) = \sum_{\omega_i \in A} P(\omega_i)$
    - $P(\Omega) = \sum_{\omega_i \in \Omega} P(\omega_i) = 1$

- 예 – [동전 던지기]
  - 1개의 공정한 동전을 세 번 던졌을 때 정확히 앞면이 한 번 나올 확률 :

    $$P(A)=\frac{3}{8}$$



  - 확률의 성질 :
    - $0 \le P(A)=\frac{3}{8} \le 1$
    - $P(A)=P(\omega_4)+P(\omega_6)+P(\omega_7)=\frac{1}{8}+\frac{1}{8}+\frac{1}{8}=\frac{3}{8}$
    - $$\begin{aligned} P(\Omega) &= P(\omega_1)+P(\omega_2)+P(\omega_3)+P(\omega_4) \\ &\quad +P(\omega_5)+P(\omega_6)+P(\omega_7)+P(\omega_8) \\ &= \frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}=1 \end{aligned}$$

확률의 법칙

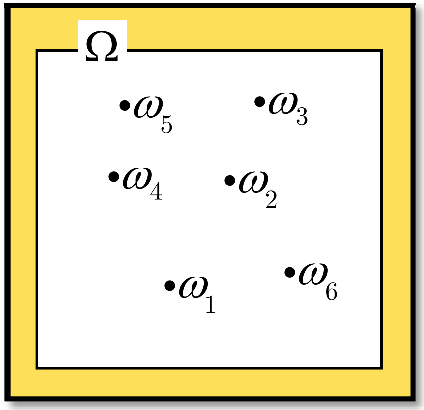
- 확률에 관한 법칙
  - 사상들의 기본적인 연산법칙
    - 여사상(complementary event)
      - 사상 A에 포함되지 않는 근원사상들의 모임
      - A^c
    - 합사상(union of events)
      - 사상 A 혹은 B에 포함되는 근원사상들의 모임
      - $A \cup B$
    - 곱사상(product events)
      - 사상 A와 B에 동시에 포함되는 근원사상들의 모임
      - $A \cap B$
    - 배반사상(disjoint event, mutually exclusive event)
      - 동시에 일어날 수 없는 사상
      - $A \cap B = \varnothing$ 인 두 사상 A, B

- 확률 기본 법칙
  - 여사상의 확률 법칙
    - 사상 A에 속하는 근원사상 + 사상 A에 속하지 않는 근원사상
      = 표본 공간 내의 모든 근원사상
    - $P(A^c) = 1 - P(A)$
  - 합사상의 확률 법칙
    - 사상 A와 B의 합사상
      = 사상 A + 사상 B − 사상 A와 B의 곱사상
    - $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
    - 사상 A와 B가 배반사상인 경우 :
      $P(A \cup B) = P(A) + P(B)$

- 예 – [원숭이의 약물 실험]
  - 실험실에서 약물실험에 쓰이는 원숭이 네 마리의 종류와 나이 조사

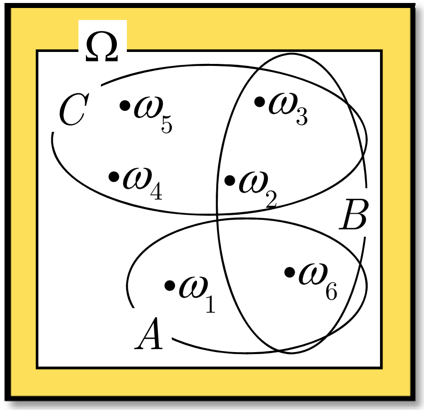
원숭이	종류	나이
1	Baboon	6
2	Baboon	8
3	Spider	6
4	Spider	6

  - 두 마리의 원숭이를 임의 선택 후 약물을 투여
    - 표본공간 :
      $\Omega = \{(1,2),\ (1,3),\ (1,4),\ (2,3),\ (2,4),\ (3,4)\}$
    - 근원사상 :
      $\omega_1 = (1,2),\ \omega_2 = (1,3),\ \omega_3 = (1,4),$
      $\omega_4 = (2,3),\ \omega_5 = (2,4),\ \omega_6 = (3,4)$



- 예 – [원숭이의 약물 실험]
  - 선택된 두 마리의 원숭이들이 같은 종류인 사상 :
    $A = \{\ \omega_1,\ \omega_6\ \} = \{(1,2),\ (3,4)\}$
  - 선택된 두 마리의 원숭이들이 같은 나이인 사상 :
    $B = \{\ \omega_2,\ \omega_3,\ \omega_6\ \} = \{(1,3),\ (1,4),\ (3,4)\}$
  - 선택된 두 마리의 원숭이들이 다른 종류인 사상 :
    $C = \{\ \omega_2,\ \omega_3,\ \omega_4,\ \omega_5\ \} = \{(1,3),\ (1,4),\ (2,3),\ (2,4)\}$

  - 사상 A 와 사상 C는 배반사상
    $A \cap C = \varnothing$
  - 선택된 두 마리의 원숭이들이 같은 종류이거나 같은 나이일 확률은?
    $$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
    $$= \frac{2}{6} + \frac{3}{6} - \frac{1}{6} = \frac{2}{3}$$



- 조건부확률과 확률법칙
  - 조건부확률(conditional probability)
    - 사상 B에 관한 정보가 주어졌을 때 사상 A가 발생할 확률

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \qquad (\text{단, } P(B) > 0)$$

  - 곱사상의 확률법칙
    - 표본공간에서 정의되는 두 개의 사상 A 와 B 에 대해 $P(A) > 0,\ P(B) > 0$ 라면,
    - $P(A \cap B) = P(B)P(A \mid B)$
      $= P(A)P(B \mid A)$

독립과 종속

- 독립사상과 종속사상
  - 독립 (independent)
    - 표본공간에서 정의되는 두 사상 A와 B의 발생이 서로 무관함
    - 독립의 조건

$$P(A \cap B) = P(A)\ P(B)$$

    - 두 사상 A와 B가 독립이면 $P(B \mid A) = P(B),\ P(A \mid B) = P(A)$
  - 종속 (dependent)
    - 두 사상 A와 B의 발생이 서로 영향을 받음
    - $$\begin{aligned} P(A \cap B) &= P(B)P(A \mid B) \\ &= P(A)P(B \mid A) \\ &\neq P(A)P(B) \end{aligned}$$

- 예 – [몸무게와 혈압]
  - 어떤 사무실 직원들을 몸무게와 고혈압의 상태에 따라 분류

	비만	정상	체중미달	계
고혈압	0.10	0.08	0.02	0.20
비고혈압	0.15	0.45	0.20	0.80
계	0.25	0.53	0.22	1.00

  - 임의의 한 직원이 고혈압일 사상 : A
  - 임의의 한 직원이 비만일 사상 : B
  - 임의로 뽑은 한 사람이 비만이었다. 그 사람이 또한 고혈압일 확률은?

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{0.10}{0.25} = 0.4$$

  - 사상 A 와 사상 B 는 서로 독립인가?

$$P(A \cap B) = 0.10 \neq P(A)\ P(B) = 0.20 \times 0.25 = 0.05$$