인공지능 수학: 통계학
2. 확률

윤상민

E-mail : smyoon@kookmin.ac.kr
Office : 02-910-4645

확률 (probability)

- 상대 도수에 의한 정의
  - 똑같은 실험을 무수히 많이 반복할 때 어떤 일이 일어나는 비율
    - 상대도수의 극한
  - 다음날 비가 올 확률?

확률 (probability)

- 고전적 정의
  - 표본공간 (sample space)
    - 모든 가능한 실험결과들의 집합
    - 예) 주사위의 숫자: {1,2,3,4,5,6}
  - 사건
    - 관심있는 실험결과들의 집합
    - 표본 공간이 부분집합
    - 예) 주사위의 숫자 중 짝수: {2,4,6}
  - 어떤 사건이 일어날 확률
    - 표본 공간의 모든 원소가 일어날 확률이 같은 경우
    - 사건의 원소의 수 / 표본공간의 원소의 수

확률 (probability)

- 고전적 정의
  - 주사위를 2번 던졌을 때 합이 10일 확률
  - 표본 공간
    - 총 36개의 원소
    - $\{(1,1), (1,2), \ldots,(1,6), (2,1), (2,2),\ldots(6,6)\}$
  - 합이 10일 사건
    - $\{(4,6), (5, 5), (6,4)\}$
    - 확률 = 3/36 = 1/12
- A: 어떤 사건
  - A가 일어날 확률
  - $P(A)$

확률 (probability)

- 확률 1
  - 반드시 그 사건이 일어남
- 확률 0
  - 그 사건이 절대로 일어나지 않음
- 확률은 0에서 1 사이의 값을 가짐

확률의 계산

- 고전적 확률
  - 표본 공간의 원소의 수를 세야함
  - 사건의 원소의 수를 세야 함
  - 따라서 경우의 수를 쉽게 셀 수 있는 방법이 필요
  - 조합 (combination) 사용

확률의 계산

- 조합 (combination)
  - 어떤 집합에서 순서에 상관없이 뽑은 원소의 집합
- n개 중 r 개를 뽑는 조합의 수

$$ {}_nC_r = \binom{n}{r} = \frac{n!}{r!\,(n-r)!} $$

$$ n! = n(n-1)(n-2)\ldots 2 \cdot 1 $$

$$ \binom{5}{2} = \frac{5!}{2!\,(3)!} = \frac{5 \times 4 \times 3 \times 2 \times 1}{2 \times 1 \times 3 \times 2 \times 1} = 10 $$

확률의 계산

- 검은공: 3개
- 흰공: 4개
- 2개의 공을 무작위로 뽑을 때, 둘다 흰공이 나올 확률?
  - 1번부터 7번까지 7개의 공
  - 1~3: 검은공
  - 4~7: 흰공
  - 표본공간
    - {(1,2), (1,3),…,(6,7)}
  - 흰공 2개인 사건
    - {(4,5), (4,6), (4,7), (5,6), (5,7), (6,7)}

확률의 계산

- 검은공: 3개
- 흰공: 4개
- 2개의 공을 무작위로 뽑을 때, 둘다 흰공이 나올 확률?
  - 표본 공간의 원소의 수
    - $\binom{7}{2} = 21$
  - 흰공이 2개 뽑히는 경우의 수
    - $\binom{4}{2} = 6$
  - 확률$= \frac{6}{21} = \frac{2}{7}$

확률의 계산

- 검은공: 3개
- 흰공: 4개
- 3개의 공을 무작위로 뽑을 때, 흰공 1개, 검은공 2개가 나올 확률?
  - 표본 공간의 원소의 수
    - $\binom{7}{3} = 35$
  - 흰공 1개, 검은공 2개 뽑히는 경우의 수
    - $\binom{4}{1} \times \binom{3}{2} = 12$
    - 두개의 경우의 수를 곱한 이유?
  - 확률 $= \frac{12}{35}$

덧셈 법칙 (Addition Law)

- 주사위를 던지는 실험
  - 표본 공간: $S = \{1,2,3,4,5,6\}$
- 사건 A
  - 주사위의 숫자가 짝수인 사건
  - $P(A) = \frac{1}{2}$
- 사건 B
  - 주사위의 숫자가 4이상인 사건
  - $P(B) = \frac{1}{2}$

덧셈 법칙 (Addition Law)

- 사건 A나 사건 B가 일어날 확률
  - $A \cup B = \{2,4,5,6\}$
  - $P(A \cup B) = \frac{|A \cup B|}{|S|} = \frac{4}{6} = \frac{2}{3}$
- 사건 A와 사건 B가 동시에 일어날 확률
  - $A \cap B = \{4,6\}$
  - $P(A \cap B) = \frac{|A \cap B|}{|S|} = \frac{2}{6} = \frac{1}{3}$
- $\boldsymbol{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$
  - $P(A \cup B) = \frac{1}{2} + \frac{1}{2} - \frac{1}{3} = \frac{2}{3}$

덧셈 법칙 (Addition Law)

- 1000명의 사람이 있는데, 남자의 비율이 40%, 20세 미만의 비율이 43%, 20세 미만이면서 남자인 사람의 비율이 15%라고 한다.
- 한 명의 사람을 랜덤하게 뽑을 때 남자이거나 20세 미만일 확률?

덧셈 법칙 (Addition Law)

- A: 남자일 사건
  - $P(A) = 0.4$
- B: 20세 미만일 사건
  - $P(B) = 0.43$
- $P(A \cap B) = 0.15$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
  - $P(A \cup B) = 0.4 + 0.43 - 0.15 = 0.68$

서로 배반 (Mutually Exclusive)

- 두 사건의 교집합이 공집합일 경우
  - 사건 A와 사건 B가 서로 배반
  - $P(A \cap B) = 0$
  - $P(A \cup B) = P(A) + P(B)$
- 사건 A
  - 주사위를 던져서 홀수가 나오는 사건
  - {1, 3, 5}
- 사건 B
  - 주사위를 던져서 4의 배수 나오는 사건
  - {4}
- A와 B는 서로 배반
  - $P(A \cup B) = P(A) + P(B) = \frac{1}{2} + \frac{1}{6} = \frac{2}{3}$

조건부 확률

- conditional probability
- 어떤 사건 A가 일어났을 때, 다른 사건 B가 일어날 확률

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

  - 단, $P(A) > 0$

조건부 확률

- 주사위를 하나 던졌는데, 4이상의 수가 나옴
- 이 때 그 수가 짝수일 확률?
  - 사건 A
    - 4이상의 수가 나오는 사건
    - $P(A) = \frac{3}{6} = \frac{1}{2}$
  - 사건 B
    - 짝수가 나오는 사건
  - $P(A \cap B) = \frac{2}{6} = \frac{1}{3}$

표본 공간의 변화

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{1/3}{1/2} = \frac{2}{3}$$

곱셈법칙

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

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

- 어떤 학교에서 60%의 학생이 남학생이다.
- 그 학교 남학생의 경우 80%는 축구를 좋아한다.
- 그 학교에서 학생 1명을 랜덤하게 뽑았을 때 축구를 좋아하는 남학생일 확률은?
  - $P(A \cap B) = P(B|A)P(A) = 0.8 \times 0.6 = 0.48$

서로 독립

- $P(B|A) = P(B)$인 경우
  - 사건 A와 B는 서로 독립
- $P(A \cap B) = P(B|A)P(A) = P(B)P(A) = P(A)P(B)$

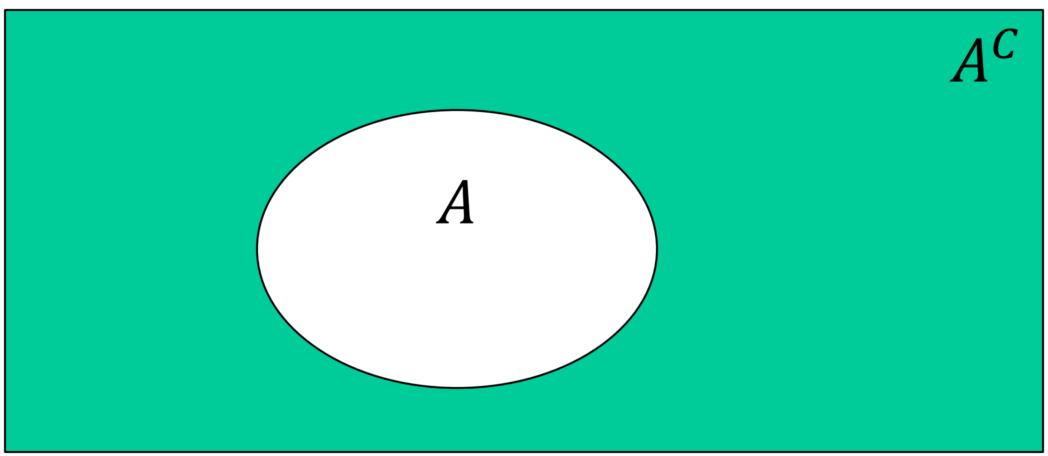
- 주사위를 2개 던지는 실험
  - 사건 A
    - 첫번째 주사위의 숫자가 짝수인 사건
  - 사건 B
    - 두번째 주사위의 숫자가 짝수인 사건
  - $P(A \cap B) = P(A)P(B) = \frac{1}{4}$

여사건

- 사건 A의 여사건
  - 사건 A가 일어나지 않을 사건
  - A^C로 표시
- 주사위 1개를 던지는 실험
  - 사건 A
    - 주사위의 숫자가 짝수인 사건
  - 사건 A의 여사건
    - 주사위의 숫자가 짝수가 아닐 사건
    - 즉, 주사위의 숫자가 홀수일 사건

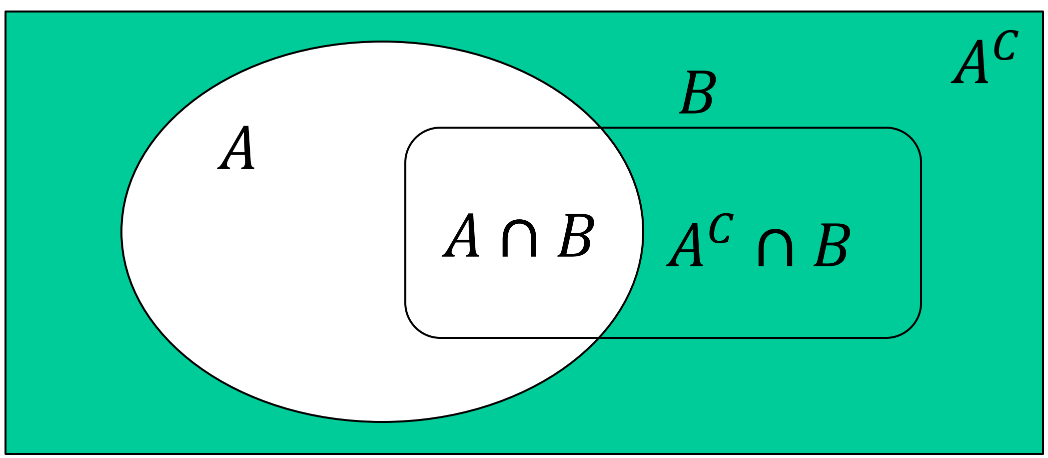
여사건

- 어떤 사건과 그 여사건은 서로 배반
  - 둘 중 하나는 반드시 일어남
  - $P(A \cup A^C) = P(A) + P(A^C) = 1$
  - $P(A) = 1 - P(A^C)$



확률의 분할 법칙

- 사건 B는 다음과 같이 나누어짐
  - $B = (A \cap B) \cup (A^C \cap B)$
  - $(A \cap B)$와 $(A^C \cap B)$는 서로 배반
  - 따라서
    - $P(B) = P[(A \cap B) \cup (A^C \cap B)] = P(A \cap B) + P(A^C \cap B)$



확률의 분할 법칙

$$P(B) = P(A \cap B) + P(A^C \cap B)$$
$$= P(B|A)P(A) + P(B|A^C)P(A^C)$$

- 어떤 사파리에서는 70%가 사자이고, 나머지가 호랑이다.
- 사자는 60%가 2살 이상이고, 호랑이는 40% 정도가 2살 이상이다.
- 전체 동물 중 2살 이상인 동물의 비율은?
  - 사건 A: 동물이 사자인 사건
  - 사건 B: 동물이 2살 이상인 사건
  - $P(B) = P(B|A)P(A) + P(B|A^C)P(A^C) = 0.6 \times 0.7 + 0.4 \times 0.3 = 0.54$

베이즈 정리

- 앞의 예제에서 동물 한마리를 랜덤하게 선택했는데, 이 동물이 2살 이상이었다.
- 이 동물이 사자일 확률은?
  - $P(A|B)$?
  - $P(A)$와는 다름.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^C)P(A^C)}$$

$$= \frac{0.6 \times 0.7}{0.6 \times 0.7 + 0.4 \times 0.3} = 0.78$$

- 아무 정보가 없었을 때 사자일 확률은 0.7이지만, 2살 이상이라는 추가 정보로 인해 그 확률이 0.78이 됨.

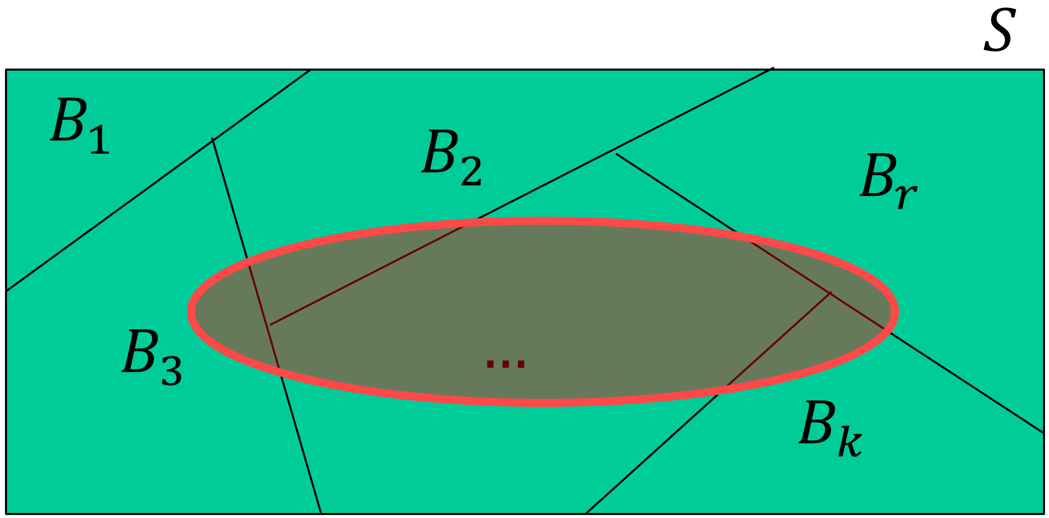
베이즈 정리

- 처음의 확률:
  - 사전확률 (prior probability)
- 수정된 확률:
  - 사후확률 (posterior probability)

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^C)P(A^C)}$$

베이즈의 정리

- 사건 $B_1, B_2, \ldots, B_k$ 가 표본 공간 S의 분할
- $P(B_r|A) = \frac{P(B_r \cap A)}{P(A)} = \frac{P(B_r \cap A)}{\sum_{i=1}^{k} P(B_i \cap A)} = \frac{P(B_r)P(A|B_r)}{\sum_{i=1}^{k} P(B_i)P(A|B_i)}$



베이즈 정리

- 어떤 사람이 검은색과 흰색의 셔츠를 가지고 있는데, 매일 아침 3/4정도는 검은색 셔츠를 입고, 1/4 정도는 흰색 셔츠를 입는다.
- 이 사람이 검은색 셔츠를 입었을 때는 3/4 정도 넥타이를 매고, 흰색 셔츠를 입었을 때는 1/2 정도 넥타이를 맨다고 하자.
- 어느날 이 사람이 넥타이를 맸다면, 이 사람이 검은색 셔츠를 입었을 확률을 구하시오.

베이즈 정리

- 사건 A
  - 아침에 검은색 셔츠를 입는 사건
  - $P(A) = \frac{3}{4}$
- 사건 B
  - 넥타이를 맨 사건
  - $P(B|A) = \frac{3}{4}, P(B|A^C) = \frac{1}{2}$
- 구하는 확률
  - $P(A|B)$?

$$P(A|B) = \frac{(3/4) \times (3/4)}{(3/4) \times (3/4) + (1/2) \times (1/4)} = \frac{9}{11}$$