2020/09/03, 2020 확률 세미나

확률 및 통계학

- 2장 확률 (Probability) -

박 재 형(jaehyoung@pel.sejong.ac.kr)

세종대학교 프로토콜공학연구실

목 차

- 표본 공간 (Sample space)
- 사상 (Event)
- 경우의 수 (Number of cases)
- 사상의 확률
- 독립사상, 종속사상
- 가법 정리, 승법 정리
- 조건부 확률
- 베이즈 정리

표본 공간 (Sample space)

- 정의
  - 통계적 실험에서 발생가능한 모든 결과들의 집합

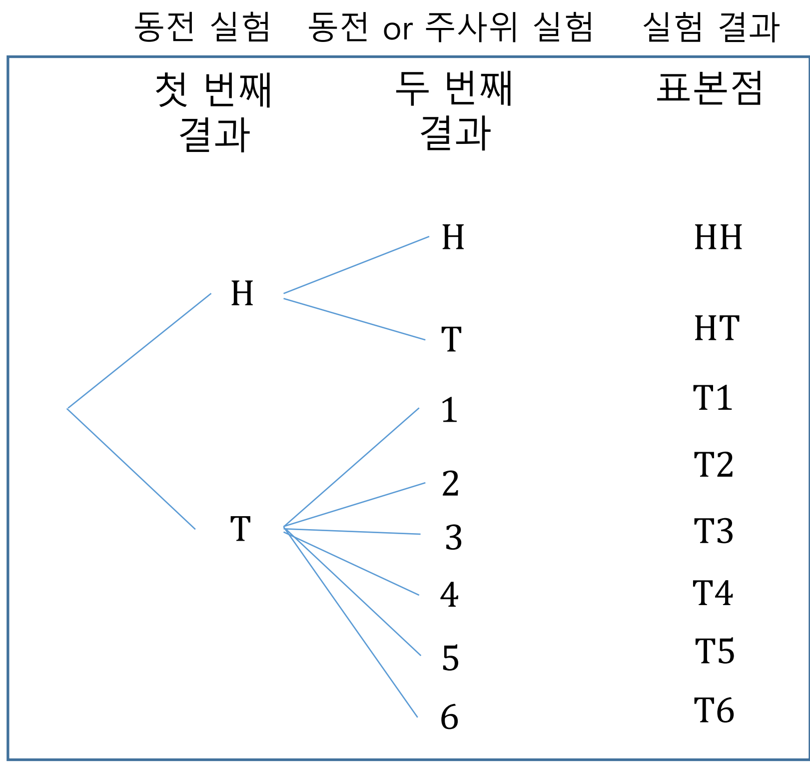
- 특징
  - 표본 공간은 표본점들로 구성 되어있음
    - 표본점: 표본의 원소 (Element 또는 Member)를 표본점 (Sample point)이라고 함

  - 기호 S로 표시
    - $S = \{$표본점 1, 표본점 2, $\cdots\}$

  - 실험마다 하나 이상의 표본공간이 사용될 수 있음
    - 표본공간은 실험의 결과의 가장 많은 정보가 포함되어야 함

표본 공간 (Sample space)

- 수형도 (Tree Diagram)
  - 표본공간을 트리구조로 표현하여 원소들을 나열할 수 있는 다이어그램
  - 수형도 표현
    - 실험 예제
      - 동전 앞면 $(H) \rightarrow$ 동전
      - 동전 뒷면 $(T) \rightarrow$ 주사위 (1~6)
    - $S = \{HH, HT, T1, T2, T3, T4, T5, T6\}$

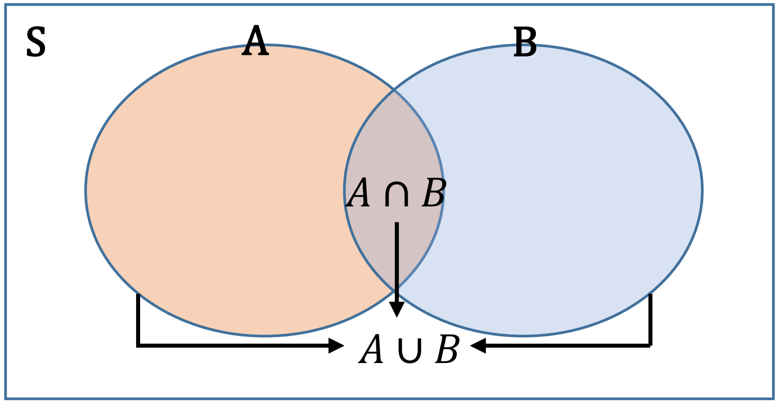


표본 공간 (Sample space)

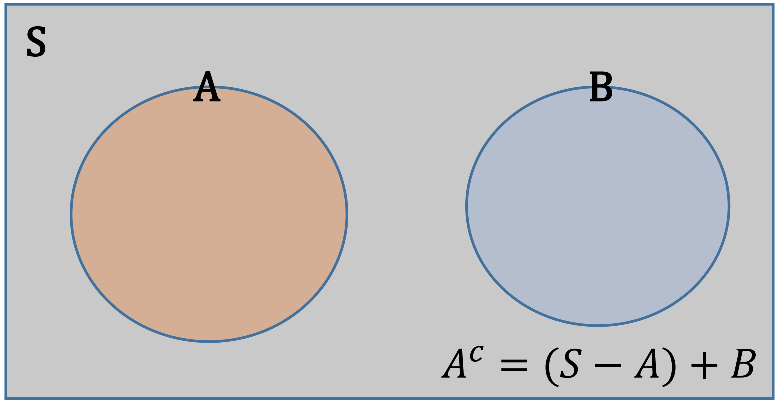
- 식 표현 예제
  - $S = \{H, T\}$
    - 동전 한 개를 던지는 실험
  - $S = \{x \mid x$는 인구 백만이 넘는 도시$\}$
    - 전세계 인구 백만이 넘는 도시들의 집합
  - $S = \{(x, y) | x^2 + y^2 \le 4\}$
    - 반지름이 2인 원의 원주상과 내부에 있는 모든 점(x, y)의 집합
  - $S = \{D, ND, NND\}$
    - 제품 3개중 불량품 D가 발견될 때까지 표본을 추출

사상 (Event)

- 정의
  - 표본공간의 부분집합을 의미
    - 공집합, 여집합이 될 수 있음
    - 사건이라고도 지칭함



- 표본공간과 사상의 관계를, 벤 다이어그램을 이용한 집합개념 표현



  - 공집합 (Empty Set): $\emptyset$
  - 여집합 (Complement): A^c
  - 교집합 (Intersection): $A \cap B$
  - 합집합 (Union): $A \cup B$
  - 상호배반 (Mutually Exclusive or Disjoint): $A \cap B = \emptyset$

사상 (Event)

- 예제
  - Q1
    - 원점에 중심을 두고 반지름이 3인 원의 내부에서 제 1사분면에 있는 점들로 구성된 표본 공간을 수식을 사용하여 표현하라
      - $S = \{(x, y) \mid 0 \leq x \ \leq 3, 0 \leq y \leq 3\}$
  - Q2
    - 다음 중 동일한 사상은 어느 것인가? (1,3)
      1. $A = \{1, 3\}$
      2. $B = \{x \mid x$는 주사위 눈금 수$\}$
      3. $C = \{x \mid x^2 - 4x + 3 = 0\}$
      4. $D = \{x \mid x$는 6개의 동전을 동시에 던졌을 때 나오는 앞면의 수$\}$

경우의 수 (Number of cases)

- 정의
  - 실험에서 발생할 수 있는 모든 사상 (Event)의 경우를 숫자로 표현한 것

- 경우의 수 계산 방법
  - 합의 법칙 $(A \cup B)$
    - 사상 A **또는** B 가 독립적으로 일어나는 경우의 수
    - $A + B$

  - 곱의 법칙 $(A \cap B)$
    - 사상 A, B가 **동시**에 일어나는 경우의 수
    - $A \times B$

경우의 수 (Number of cases)

- 순열 (Permutation)
  - 정의
    - 특정 집합의 전체 또는 일부의 순서적 배열
      - 식표현: $n! = n(n-1)\cdots(2)(1)$
        - e.g., (a), (b), (c), (d)의 배열
    - 서로 다른 n개의 원소 중 r개를 중복없이 골라 순서에 상관 있게 나열하는 경우의 수
      - 식표현: $nPr = \frac{n!}{(n-r)!}$
        - e.g., (a), (b), (c), (d)에서 중복 없이 고른 2개의 배열
  - 원순열 (Circular Permutation)
    - 서로 다른 n개를 한 원소를 기준으로 원형으로 나열하는 경우의 수
    - 식표현: $(n-1)!$
      - e.g., 원형 테이블에서 4명의 사람이 카드놀이를 하는 경우 자리배치

경우의 수 (Number of cases)

- 순열 (Permutation)
  - 분할 (Partition)
    - n개의 서로 다른 대상물을 r개의 부분집합으로 나열하는 경우의 수
    - 식표현: $\binom{n}{n_1!,\ n_2!,\ \cdots\ n_r} = \frac{n!}{n_1!n_2!\cdots n_r}$
      - e.g., $\{a, e, i, o, u\}$에서 한 쪽에4개의 원소를 포함하고 다른 한쪽에 1개의 원소를 포함하는 경우
  - 조합 (Combination)
    - n개의 대상물 중에서 r개의 대상물을 순서를 고려하지 않고 선택하는 경우의 수
    - 식표현: $nCr = \binom{n}{r, n-r} = \binom{n}{r} = \frac{n!}{(n-r)!r!} = \frac{nPr}{r!}$
      - e.g., 5개의 게임 중 2개의 스포츠 게임을 고르는 경우

독립 사상, 종속 사상

- 독립 사상 (Independence Event)
  - 정의
    - 서로 다른 사건 A, B에 대하여 A와 B의 확률이 상호 관계가 없는 경우 (복원 추출)
      - 복원추출: 추출한 것을 원래대로 돌려놓고 다시 추출하는 것
  - 만족하는 조건
    - $P(B|A) = P(B)$
    - $P(A|B) = P(A)$
    - $P(A \cap B) = P(A)P(B)$

독립 사상, 종속 사상

- 종속 사상 (Dependent Event)
  - 정의
    - 서로 다른 사건 A, B에 대하여 A와 B의 확률이 상호 관계가 있는 경우 (비복원 추출)
      - 비복원 추출: 추출한 것을 원래대로 돌려놓지 않고 추출하는 것
  - 만족하는 조건
    - $P(B|A) = \frac{P(B\cap A)}{P(A)}$
    - $P(A|B) = \frac{P(A\cap B)}{P(B)}$
    - $P(A \cap B) = P(B|A) \times P(A)$ 또는 $P(A|B) \times P(B)$

사상의 확률

- 정의
  - 실험에서 사상에 대한 확률 (Probability)을 의미
- 특징
  - 사상 A가 발생할 가능성을 나타내는 0~1 사이의 값
    - 어떤 사건이 일어날 확률(P)의 범위
  - 표본 공간의 모든 표본점 확률의 합
    - $P(S) = 1$
  - 상호배반 사상의 경우 각 사상의 확률의 합
    - $P(E_1 + E_2 + E_3) = P(E_1) + P(E_2) + P(E_3)$

가법 정리, 승법 정리

- 가법 정리 (Addition Theorem)
  - 정의
    - A, B 중 적어도 하나가 일어날 확률을 계산하는 정리
  - 확률의 덧셈정리
    - $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
    - 두 사상 A, B가 서로 상호배반 사상일 경우, $A \cap B = \emptyset$이기 때문에 $P(A \cup B) = P(A) + P(B)$
  - 여집합 관계
    - 여집합: 사상 A 가 일어나지 않을 확률
    - $P(A) = 1 - P(A')$

가법 정리, 승법 정리

- 가법 정리 예제
  - Q1
    - 회사 A에 합격할 확률은 0.8 회사 B에 합격할 확률은 0.6, 두 회사 모두 합격할 경우의 확률은 0.5일 경우, 두 회사중 적어도 하나의 회사에 합격할 확률은?
      - $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.8 + 0.6 - 0.5 = 0.9 = 90\%$
  - Q2
    - 두개의 주사위를 던졌을 때, 나타난 두 눈의 합이 7이나 11이 될 확률은 얼마인가?
      - $P(A)$ =합이 7인 사상
      - $P(B)$ =합이 11인 사상
      - $P(A \cup B) = P(A) + P(B) = \frac{1}{6} + \frac{1}{18} = \frac{2}{9} =$ 약 22.2%

가법 정리, 승법 정리

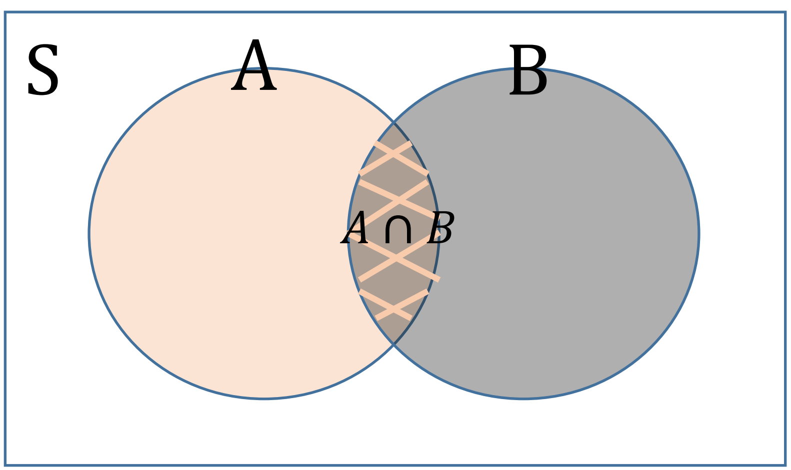
- 승법 정리 (Multiplicative Rule)
  - 정의
    - 조건부 확률 공식에 의해서 도출되며 두 사건이 동시에 일어날 확률을 계산하는 정리
  - 확률의 곱셈 정리
    - 종속 사상일 경우 (비복원 추출)
      - $P(A \cap B) = P(A)P(B|A)$ or $P(B)P(A|B)$
    - 독립 사상일 경우 (복원 추출)
      - $P(A \cap B) = P(A)P(B)$

가법 정리, 승법 정리

- 승법 정리 예제
  - Q1
    - 퓨즈 20개중 5개가 불량품이고, 2번 추출할 때, 둘 다 불량품일 확률은?
      - 독립 사상일 경우 (복원 추출): $\frac{5}{20} \times \frac{5}{20} = \frac{1}{16} = 6.25\%$
      - 종속 사상일 경우 (비복원 추출): $\frac{5}{20} \times \frac{4}{19} = \frac{1}{19} = \text{약 } 5.2\%$

조건부 확률

- 조건부 확률 (Conditional Probability)
  - 정의
    - 사상 A가 일어날 경우 사상 B가 나타날 확률
  - 표기법: $P(B|A)$
  - 식: $P(B|A) = \frac{P(B \cap A)}{P(A)}, P(A) > 0$
  - 벤 다이어그램



조건부 확률

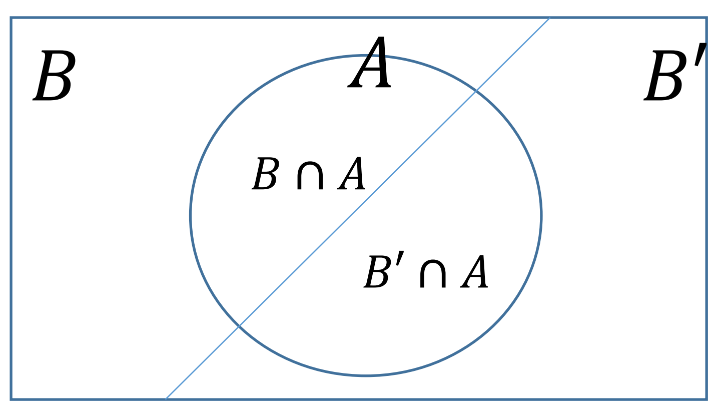
- 조건부 확률 (Conditional Probability)
  - 예시
    - 200명 대상으로 성별 스포츠 취미 조사

교육 정도	남	여
축구	38	45
농구	28	50
야구	22	17

- 축구를 좋아하는 사람이 뽑혔을 때 남자일 확률
  - $\frac{\text{축구를 좋아하는 남자}}{\text{축구를 좋아하는사람}} = \frac{38}{38+45} = \frac{38}{83} = \text{약 } 45.7\%$
- 여자가 뽑혔을 때 그녀가 농구를 좋아할 확률
  - $\frac{\text{농구를 좋아하는 여자}}{\text{여자}} = \frac{50}{112} = \frac{25}{56} = \text{약 } 44.6\%$

베이즈 정리

- 전확률의 법칙 (Law of Total Probability)
  - 정의
    - 표본 공간 S를 연속된 실험 n개의 영역으로 나눴을 때, S의 사건 A의 확률을 계산하는 법칙
      - 2개의 영역으로 나눈 경우 $(n = 2)$
        - $P(A) = P[(B \cap A) \cup (B' \cap A)]$ 가법 정리
          $= P(B \cap A) + P(B' \cap A)$ 승법 정리
          $= P(B)P(A|B) + P(B')P(A|B')$
      - 벤 다이어그램
        - $P(B \cap A),\ P(B' \cap A)$는 서로 배반 사상

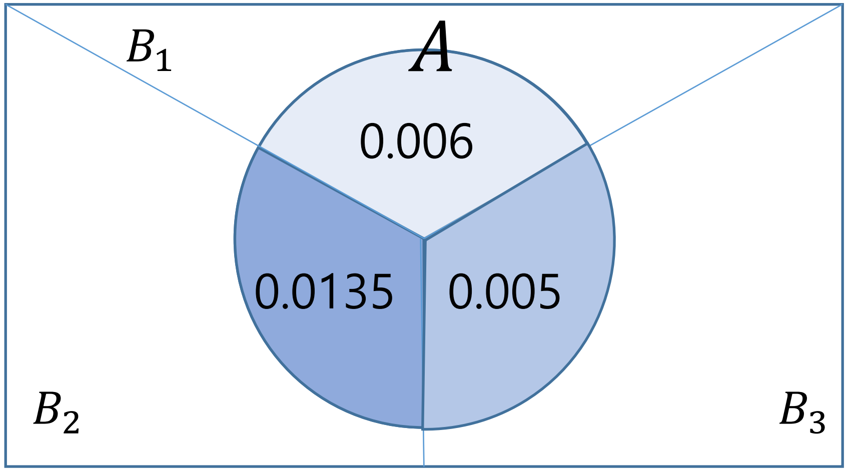


베이즈 정리

- 전확률의 정리 (Theorem of Total Probability)
  - 식을 일반화 하여 공식을 도출
  - 식표현: $P(A) = \sum_{i=1}^{k} P(B_i \cap A) = \sum_{i=1}^{k} P(B_i)P(A|B_i)$
  - 조건
    - 사건$B_1, B_2, \cdots, B_k$를 표본공간 S의 분할이고, $P(B) \neq 0, i = 1, 2, 3, \cdots, k$이면 S의 임의 사건 A는 위 식을 성립

베이즈 정리

- 전확률의 법칙 (Law of Total Probability)
  - 예제
    - Q1
      - 3대의 기계 B_1, B_2, B_3가 각각 전체 생산량의 30%, 45%, 25% 를 생산하는 조립공장에서 불량품 제조율이 각각 2%, 3%, 2%이면, 불량품을 선택할 확률은 얼마인가?
        - $P(B_1)P(A|B_1) = (0.3)(0.02) = 0.006 = 0.6\%$
        - $P(B_2)P(A|B_2) = (0.45)(0.03) = 0.0135 = 1.35\%$
        - $P(B_3)P(A|B_3) = (0.25)(0.02) = 0.005 = 0.5\%$
        - $P(A) = 0.006 + 0.0135 + 0.005 = 0.0245 = 2.45\%$



베이즈 정리

- 정의
  - 사전 확률 (Prior probability)에서 발생한 사건에 대한 사후 확률 (Posterior probability)을 구하는 이론
    - 사전 확률: 현재 가진 정보를 기초로 정한 초기 확률
    - 사후 확률: 추가된 정보로 부터 사전 정보를 수정한 확률

  - 식표현
    - $P(B_r)$ = 연속된 실험
    - $P(A)$ = 추가된 사건
    - $P(B_r|A) = \frac{P(B_r \cap A)}{P(A)} = \frac{P(B_r \cap A)}{\sum_{i=1}^{k} P(B_i \cap A)} = \frac{P(B_r)P(A|B_r)}{\sum_{i=1}^{k} P(B_i)P(A|B_i)}$

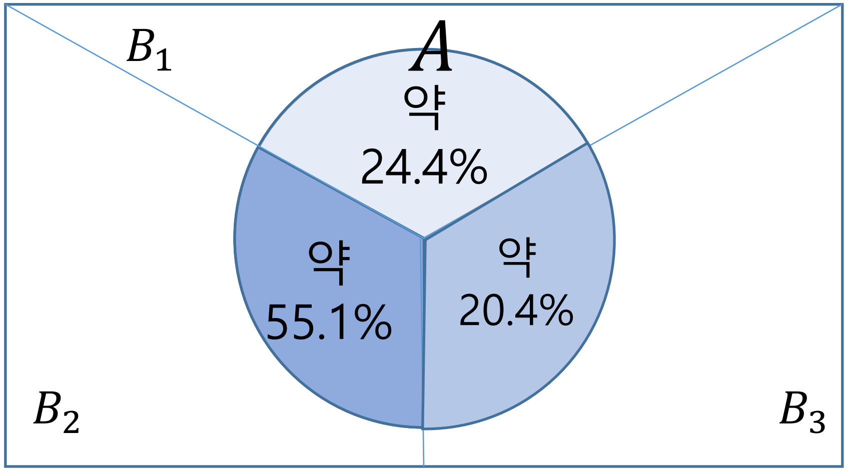
승법 정리

전 확률의 정리

베이즈 정리

- 예제
  - Q1
    - 3대의 기계 B_1, B_2, B_3 가 각각 전체 생산량의 30%, 45%, 25%를 생산하는 조립공장에서 불량품 제조율이 각각 2%, 3%, 2%이면, 불량품이 B_3 에서 제조되었을 확률은?
      - $P(B_3|A) = \frac{P(B_3)P(A \mid B_3)}{P(B_1)P(A \mid B_1)+P(B_2)P(A \mid B_2)+P(B_3)P(A \mid B_3)}$

$$= \frac{0.005}{0.006+0.0135+0.005} = \frac{0.005}{0.0245} = \frac{10}{49} = \text{약 } 20.4\%$$



Thanks!

박 재 형 (jaehyoung@pel.sejong.ac.kr)