확률 및 통계학

- 2장 확률 (Probability) -

박 재 형(jaehyoung@pel.sejong.ac.kr)

세종대학교 프로토콜공학연구실

목 차

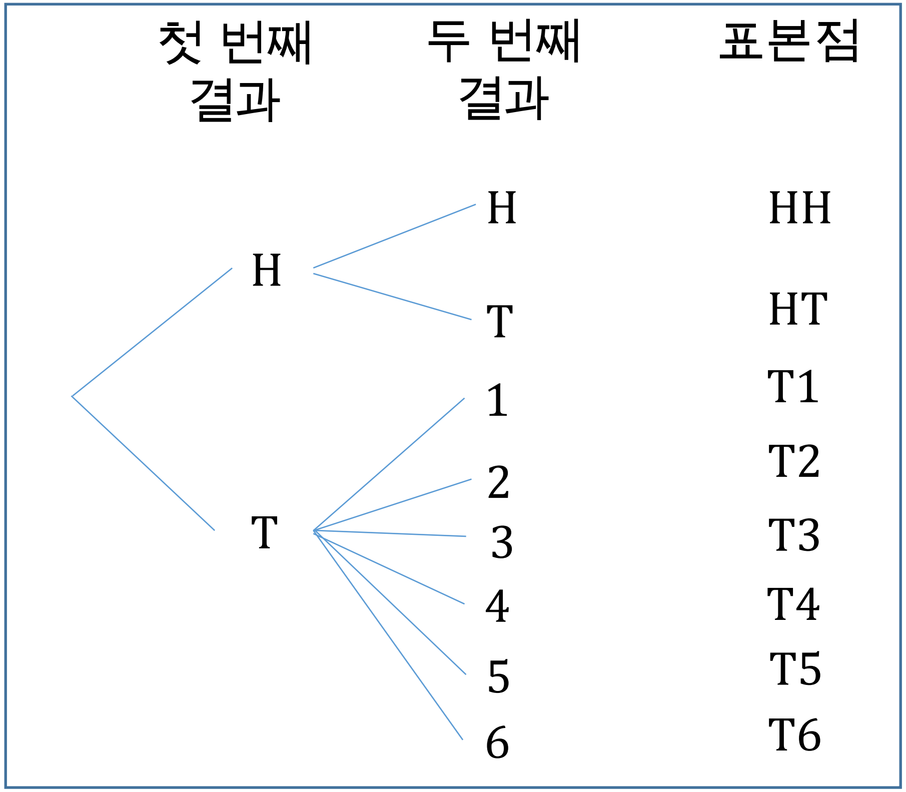
- 표본 공간 (Sample space)
- 사상 (Event)
- 경우의 수 (Number of cases)
- 사상의 확률
- 가법 정리
- 조건부 확률, 독립 사상, 승법정리
- 베이즈 정리

표본 공간 (Sample space)

- 정의
  - 통계적 실험에서 발생가능한 모든 결과들의 집합
- 특징
  - 표본 개개의 원소(Element 또는 Member)를 표본점 (Sample point)라고 함
  - 기호 S로 표시
    - $S =$ {표본점1, 표본점2}
  - 실험마다 하나 이상의 표본공간이 사용될 수 있음
    - 표본공간들 중에서도 실험의 결과로써 가장 많은 정보가 부각될 수 있는 표본공간을 사용해야 함

표본 공간 (Sample space)

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



표본 공간 (Sample space)

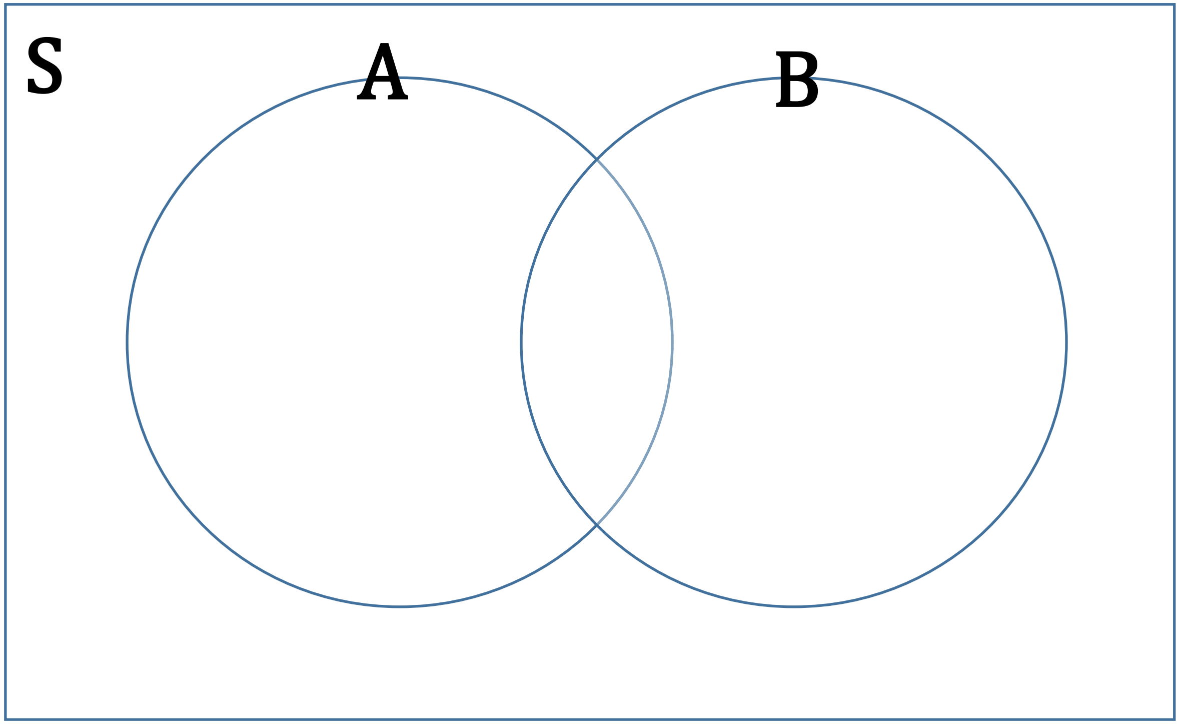
- 식 표현
  - $S = \{H, T\}$
    - 동전 한 개를 던지는 실험
  - $S = \{x \mid x$는 인구 백만이 넘는 도시$\}$
    - 전세계 인구 백만이 넘는 도시들의 집합
  - $S = \{(x, y) | x^2 + y^2 \ \leq 4\}$
    - 반지름이 2인 원의 원주상과 내부에 있는 모든 점(x, y)의 집합
  - $S = \{D, ND, NND, NNND, \cdots\}$
    - 불량품 D가 발견될 때까지 표본을 추출

사상 (Event)

- 정의
  - 표본공간의 부분집합을 의미
    - 공집합, 여집합이 될 수 있음
    - 사건이라고도 지칭함
- 표본공간과 사상의 관계를, 벤 다이어그램을 이용한 집합개념의 그래프로 표현
  - 공집합 (Empty Set)
  - 여집합 (Complement)
  - 교집합 (Intersection)
  - 합집합 (Union)
  - 상호배반 (Mutually Exclusive or Disjoint)

사상 (Event)

- 벤 다이어그램
  - $\emptyset$
  - A^c
  - $A \cap B$
  - $A \cup B$
  - $A \cap B = \emptyset$



사상 (Event)

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

경우의 수 (Number of cases)

- 정의
  - 실험에서 발생할 수 있는 모든 사상 (Event)의 가짓수
  - 표본공간의 표본점의 수를 계산하는 것
- 경우의 수 계산 방법
  - 합의 법칙 $(A \cup B)$
    - 사상 A **또는** B 가 독립적으로 일어나는 경우의 수
  - 곱의 법칙 $(A \cap B)$
    - 사상 A,B가 **동시**에 일어나는 경우의 수

경우의 수 (Number of cases)

- 순열 (Permutation)
  - 정의
    - 어떤 대상물 집합의 전체 또는 일부의 순서적 배열
      - $n! = n(n-1)\cdots(2)(1)$
      - e.g., (a), (b), (c), (d)의 배열
  - 서로 다른 n개의 원소 중 r개를 중복없이 골라 순서에 상관 있게 나열하는 것
    - $nPr = \frac{n!}{(n-r)!}$
    - 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
    - $(n-1)!$
    - e.g., 원형 테이블에서 4명의 사람이 카드놀이를 하는 경우

경우의 수 (Number of cases)

- 순열 (Permutation)
  - 분할 (Partition)
    - n개의 서로 다른 대상물을 r개의 부분집합으로 나열하는 경우의 수
    - $\binom{n}{n_1!,\ n_2!,\ \cdots\ n_r} = \frac{n!}{n_1! n_2! \cdots n_r}$
    - e.g., $\{a, e, i, o, u\}$에서 한 쪽에4개의 원소를 포함하고 다른 한 쪽에 1개의 원소를 포함하는 경우

  - 조합 (Combination)
    - n개의 대상물 중에서 r개의 대상물을 순서를 고려하지 않고 선택하는 경우의 수
    - $nCr = \binom{n}{r, n-r} = \binom{n}{r} = \frac{n!}{(n-r)!r!} = \frac{nPr}{r!}$
    - e.g., 5개의 게임 중 2개의 스포츠 게임을 고르는 경우

사상의 확률

- 정의
  - 실험에서 사상에 대한 가중치(Weight) 또는 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 각사상의 확률의 합
      - $P(E_1 + E_2 + E_3) = P(E_1) + P(E_2) + P(E_3)$

가법정리

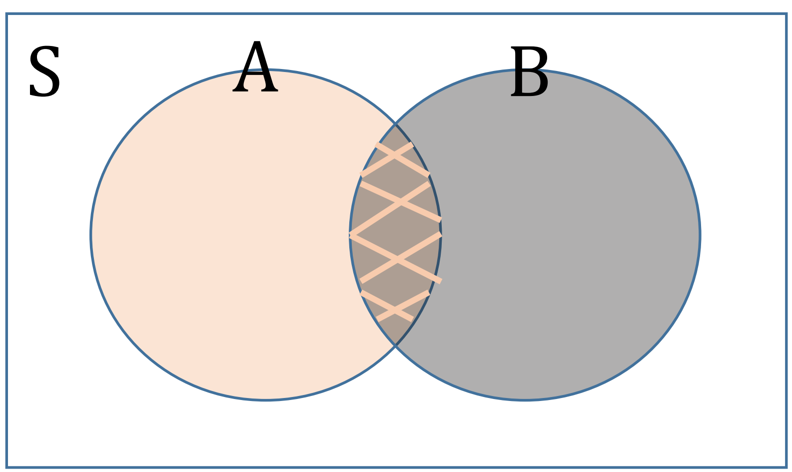
- 정의
  - 확률의 계산을 간단하게 해주는 법칙

- 확률의 덧셈정리
  - $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
  - 두 사상 A, B가 서로 상호배반 사상이면 $A \cap B = \varnothing$ 이면 $P(A \cup B) = P(A) + P(B)$

- 여집합 관계
  - 여집합: 사상 A 가 일어나지 않을 확률
  - $P(A) = 1 - P(A')$

조건부 확률, 독립 사상, 승법 정리

- 조건부 확률 (Conditional Probability)
  - 정의
    - 사상 A가 일어날 때 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, 독립 사상, 승법 정리

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

교육 정도	남	여
초등 교육	38	45
중등 교육	28	50
대학	22	17

    - 중등 교육을 마친 사람이 뽑혔을 때 남자일 확률
    - 여자가 뽑혔을 때 그녀가 대학 교육을 받지 못한 사람일 확률

조건부 확률, 독립 사상, 승법 정리

- 독립 사상 (Independence Event)
  - 정의
    - 서로 다른 사건 A, B에 대하여 A와 B의 확률이 상호 관계가 없는 경우
  - 3 조건 (하나만 만족하면 독립 사상)
    - $P(A \cap B) = P(A) * P(B)$
    - $P(B|A) = P(B)$
    - $P(A|B) = P(A)$

- 종속 사상 (Dependent Event)
  - 정의
    - 서로 다른 사건 A, B에 대하여 A와 B의 확률이 상호 관계가 있는 경우

조건부 확률, 독립 사상, 승법 정리

- 승법 정리 (Multiplicative Rule)
  - 조건부 확률 공식에 대하여 확률계산특성을 이용해 두 사건이 함께 발생할 확률을 구함
    - 두 사건 A, B가 동시에 발생할 수 있으면,
      $$P(A \cap B) = P(A)P(B|A)$$
      $$\therefore P(A \cap B) = P(B \cap A)$$
    - 어떤 사상이 A 또는 B로 표현되는지는 결과에 상관없음
  - 예제
    - 퓨즈 20개중 5개가 불량품이고, 비복원으로 2번 추출할 때, 둘 다 불량품일 확률은?

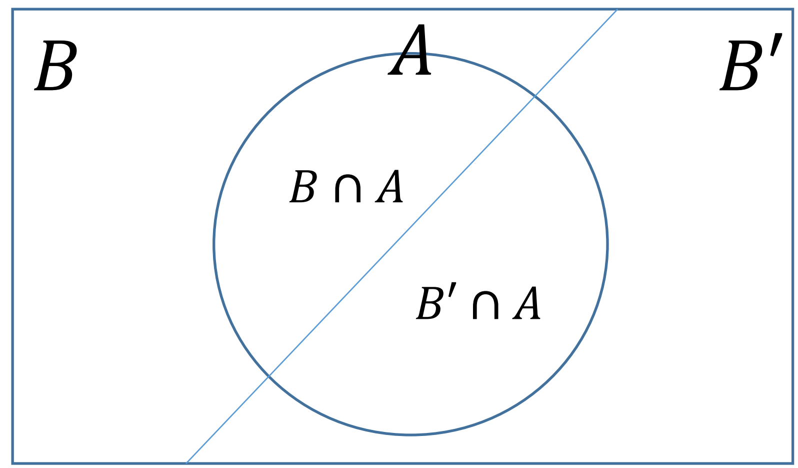
베이즈 정리

- 전확률 (Total Probability)의 법칙
  - 사상 A를 서로 배반 사상인 E와의 합집합으로 내타내는 법칙
    - $A = (B \cap A) \cup (B' \cap A)$
    - $P(A) = P[(B \cap A) \cup (B' \cap A)] = P(B \cap A) + P(B' \cap A)$
      $= P(B)P(A|B) + P(B')P(A|B')$

- 전확률의 정리 (Theorem of total probability)
  - 표본공간이 k개의 부분집합으로 분할되는 경우, 식을 일반화 하여 공식을 도출
  - $P(A) = \sum_{i=1}^{k} P(B_i \cap A) = \sum_{i=1}^{k} P(B_i)P(A|B_i)$
    - 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

- 전확률 (Total Probability)의 법칙
  - 벤 다이어그램



  - 예제
    - 3대의 기계 $B1$, $B2$, $B3$가 각각 전체 생산량의 30%, 45%, 25% 를 생산하는 조립공장에서 불량품 제조율이 각각 2%, 3%, 2%이면, 불량품을 선택할 확률은 얼마인가?

베이즈 정리

- 정의
  - 사전확률(Prior probability)과 조건부 확률로 사후확률(Posterior probability)을 구하는 이론
    - 사전 확률: 현재 가진 정보를 기초로 정한 초기 확률
      - $P(A), P(B)$
    - 조건부확률
      - $P(B|A)$
    - 사후확률: 추가된 정보로 부터 사전 정보를 수정한 확률
      - $P(A|B)$

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

베이즈 정리

- 예제
  - 40세 이상인 성인이 암을 가질 확률: 0.05
    의사가 암환자를 암이라고 진단할 확률: 0.78
    암환자가 아닌 사람을 암이라고 오진할 확률: 0.06
    - 40세 이상인 사람이 암으로 진단될 확률은?
    - 진단받은 암 환자가 실제 암 환자일 확률은?

Thanks!

박 재 형 (jaehyoung@pel.sejong.ac.kr)