■ 확률(Probability)

- 확률실험(Random experiment)
  - 실험을 하기 전까지 이들 결과 중 어떤 것이 발생할 것인지에 대해 확실하게 예측할 수 없음
  - 표본공간(sample space): 확률실험에서 발생 가능한 모든 결과들의 집합
  - 사건(event): 표본공간 내에서의 관심 부분집합
- 확률(probability): 어떤 사건이 발생할 가능성이 얼마나 되는지를 나타내는 [0,1]의 수치적 척도

■ 확률의 해석

- 고전적 확률: 표본공간의 각 원소(근원사건)의 발생가능성이 동일(equally likely)할 때 사건의 원소개수/표본공간의 원소개수
  - 경우의 수: 곱의 법칙

배열 \ 추출	복원	비복원
순서고려	Ⓐ n^k	Ⓑ $\dfrac{n!}{(n-k)!}$
순서무시	Ⓒ $\dbinom{n+k-1}{k}$	Ⓓ $\dbinom{n}{k}$

- 상대도수의 극한의 개념(통계적 확률): $P(A) = \lim_{n \to \infty} \frac{n(A)}{n}$
  - 각각의 실험에서 발생하는 결과는 표본이고 실험을 무한히 반복한다는 것은 표본이 결국 모집단이 됨
- 확률의 공리(Axiom)

【공리 1】 $P(\Omega) = 1$

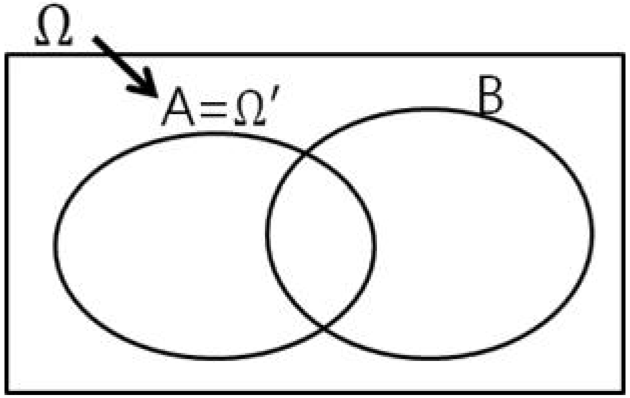
【공리 2】 $0 \le P(A) \le 1, \quad A \subset \Omega$

【공리 3】 서로 배반인 사건 $A_1, A_2, \ldots, A_n$ 에 대해

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

■ 조건부 확률(Conditional probability)

- 확률실험에서 새로운 조건(정보)이 추가되었을 때, 사건의 확률



- 사건 A 가 주어졌을 때 사건 B의 조건부 확률

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \quad P(A) > 0$$

- 조건부확률의 활용
    - $P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$
    - $$\begin{aligned} P(B) &= P(A \cap B) + P(A^c \cap B) \\ &= P(A)P(B|A) + P(A^c)P(B|A) \end{aligned}$$
    - 베이즈정리

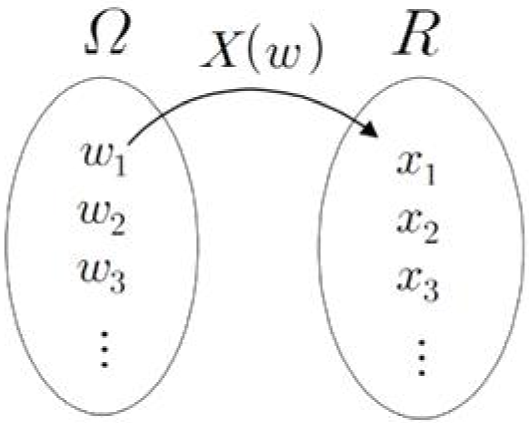
$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)\,P(B|A)}{P(A)\,P(B|A) + P(A^c)\,P(B|A^c)}$$

- "사건 A 와 B 는 독립사건(independent events)"

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

■ 확률변수(random variable)

- 표본공간에서 정의된 실함수(real-valued function)
  - ○ 정의역이 표본공간이고 공역이 실수인 함수



- 확률분포(probability distribution): 확률변수의 값에 대해 확률을 표시한 것

■ 확률분포

- 확률질량함수: 이산확률변수 X가 임의의 값 x일 확률

$$f(x) = P(X = x)$$

- 누적분포함수(cumulative distribution function)

$$P(X \leq x) = \sum_{x_i \leq x} f(x_i) \equiv F(x), \quad -\infty < x < \infty$$

- 확률밀도함수: 연속확률변수 X가 임의의 구간 $[a, b]$에 속할 확률

$$P(a \leq X \leq b) = \int_a^b f(x)\,dx$$

  - 모든 x에 대해 $P(X = x) = 0$

- 확률변수의 기댓값(expected value)
  - 확률변수에 대해 평균적으로 기대하는 값= 모평균
  - $E(X) = \sum_x x f(x) = \mu$
  - $E(X) = \int x f(x)\, dx = \mu$
- 확률변수 X 의 함수 $Y = g(X)$의 기댓값
  - $E(Y) = E(g(X)) = \sum_x g(x) f_X(x)$
  - $E(Y) = E(g(X)) = \int g(x) f_X(x)\, dx$
  - $g(X) = (X - \mu)^2$ 의 기댓값: 모분산

- 결합분포(joint distribution)
  - 두 개 이상의 확률변수들을 동시에 고려한 확률분포
  - 두 이산확률변수 X와 Y에 대해,

$$f(x,\ y) = P(X = x,\ Y = y)$$

- 주변분포(marginal distribution)

$$f_X(x) = \sum_y f(x, y), \quad f_Y(y) = \sum_x f(x, y)$$

- 두 확률변수 X와 Y는 독립 $\leftrightarrows$ 모든 x, y에 대해,

$$f(x,\ y) = f_X(x)\, f_Y(y)$$

- 확률변수 X와 Y에 대해, $X+Y$의 기댓값? XY의 기댓값?
  - 결합확률질량함수나 결합확률밀도함수를 이용
- 두 확률변수 X와 Y의 공분산

$$Cov(X, Y) = \sum_x \sum_y (x-\mu_X)(y-\mu_Y)f(x,y) = E((X-\mu_X)(Y-\mu_Y))$$

$$= \sum_x \sum_y xyf(x,y) - \mu_X\mu_Y = E(XY) - E(X)E(Y)$$

  - X와 Y가 독립이면, $E(XY) = E(X)E(Y) \Rightarrow Cov(X, Y) = 0$
- 두 확률변수 X와 Y의 상관계수

$$\rho_{XY} = Cor(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$$

확률(probability)　통계적 추론(statistical inference)

⇓ ⇓ ⇓ ⇓

- 확률: 모집단에 대해 알고 있다고 가정 하에서 표본의 성질에 대해 알아보는 것
- 통계적 추론: 수집된 표본의 정보를 이용하여 미지의 모집단을 유도하는 과정