6.2 확률의 수학적 정의와 의미

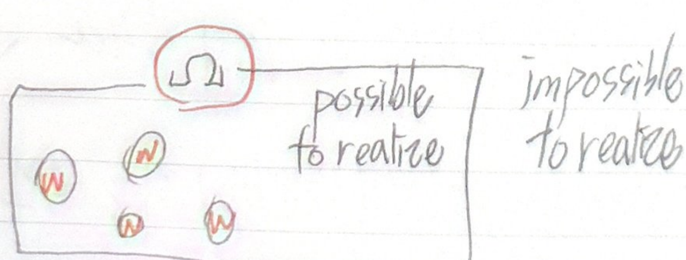
일반적으로 확률을 다룰때는 현실에서 해결하고자 하는 문제와 결부하여 정의한다. ...
확률의 수학적 정의에는 3가지 개념이 필요하다.

- 확률표본
- 표본공간
- 사건

[표본공간과 확률표본]

- 확률표본 (probablistic sample, random sample) / 표본 (Sample)
  - 풀고자 하는 확률적문제에서 발생(realize) 할수있는
    현상 또는 선택(sampled) 될 수 있는 하나의경우.

- 표본공간 (Sample space)
  - 가능한 모든 표본들의 집합.
  - Ω (Omega) 로 표시 (표본 = ω)
  - 표본공간을 정의한다 = 우리가 고려하는 범위에서 어떤 표본(경우, 현상)
    이 가능 하고 어떤 표본이 가능하지 않은지 정의.



사건 event

표본공간 Ω의 부분집합

— 전체 사건 중 우리가 관심을 갖는 일부 sample subset.

확률 probability

사건을 입력하면 숫자가 출력되는 함수

$$\text{event (subset)} \longrightarrow \boxed{\text{prob. func.}} \longrightarrow \text{number (prob. value)}$$

- 확률이라는 함수의 입력값 집합 = 정의역(domain)

  = 표본공간의 모든 사건 (부분집합) 의 집합.

- 모든 각각의 사건 (부분집합)에 어떤 숫자를 할당(assign, allocate) 하는 함수

  = 확률 (probability)

(notation) $\mathbf{P}$, $P(A)$

규칙 ① 모든 사건에 대해 확률은 실수이고 0 이상의 양수다.

$$P(A) \geq 0$$

규칙 ② 표본공간 (전체집합) 이라는 사건 (부분집합) 확률 = 1

$$P(\Omega) = 1$$

규칙 ③ 공통 원소가 없는 두 사건의 합집합의 확률은 사건별 확률의 합이다.

$$A \cap B = \emptyset \longrightarrow P(A \cup B) = P(A) + P(B)$$

콜모고로프의 공리 Kolmogorov's axioms

표본공간 Ω

Ⓗ Ⓣ

→

사건(부분집합) 집합

$A\{\}$ 0.0 $C\{T\}$ 0.4

$B\{H\}$ 0.6 $D\{H,T\}$ 1.0

[확률은 표본이 아닌 사건을 input으로 갖는 function]

"확률에 대한 오해" ⇒ 확률은 표본에 대하여 정의된 숫자!

ex 주사위 1 나타내는 경우,

$P(1) = \frac{1}{6}$ → WRONG!

올바른 식

$$\boxed{P(\{1\}) = \frac{1}{6}}$$

[육면체 주사위의 확률?]

$P(\{1\}) = \frac{1}{6}$ if pair dice*

domain 정보가 주어지지 않았다면, $P(A) = \frac{card(A)}{card(\Omega)}$ 을 각 사건의 P로 보는것이 타당하다.

공정한 주사위가 아니라면 $P(\{1\}) \neq \frac{1}{6}$ 일 수도 있다.

연습문제 6.2.7

(1) 약속날짜가 31일인지 아닌지를 결정하는 문제에 확률을 할당해보자

$\Omega = \{Y, N\}$

$P(\{\}) = 0.0$

$P(\{Y\}) = 0.95$

$P(\{N\}) = 0.05$

$P(\{Y, N\}) = 1.0$

(2) 사과와 오렌지만 파는 과일가게에서 손님이 선택한 어떤 과일인지 결정하는 문제에 대해 확률을 할당해보자.

$\Omega = [apple, orange]$

$P(\{\}) = 0.0$

$P(\{apple\}) = 0.17$

$P(\{orange\}) = 0.83$

$P(\{apple, orange\}) = 1.0$

~~(3) 사과와 오렌지~~

NO.

DATE.

확률의 의미

1. 빈도주의 (Frequentist) 관점

- 반복적으로 선택된 표본이 사건(부분집합) A의 원소가 될 경향(propensity)을 그 사건의 확률이라고 봄
- 실제 동전을 반복하여 던졌을 때, 동전을 던진 전체 횟수에 확률값을 곱한 숫자만큼 해당 사건이 발생한다고 봄.
- 10,000번 동전을 던지면 10,000 X 0.5 = 5,000번 앞면이 나온다

2. 베이지안 (Bayesian) 관점

- 선택된 표본이 특정한 사건(부분집합)에 속한다는 가설(hypothesis), 명제(proposition) 혹은 주장(assertion)의 * 신뢰도 (degree of belief) 라고 봄
- 반복이라는 개념은 사용되지 않는다

관점은 다르지만
어느 정도의 불확실성을 전제하고 있다.

(Bayesian example)

새는 날 수 있다 → 새가 날 수 있는 가능성은 95%이다

주사위 굴리면 $P(H) = 0.5$ → ~~P(H)=0.5라는 주장/가설/명제의 신뢰도가~~ H라는 사건이 발생했다는 주장/명제/가설의 신뢰도는 0.5이다.

Bayesian PoV

- 베이지안 관점에서 사건(부분집합)의 의미?

→ "원하는 답(표본)이 포함되어 있을 가능성이 있는 후보의 집합"

- 어떤 사건을 제시한다는 것의 의미

→ "이 사건에 속한 원소중에 원하는 답(표본)이 있다"는 명제/가설/주장을 제시하는 것 ex) $P(\{1,3\}) = \frac{1}{2}$ → 주장

- 사건이 일어났다(occur), 발생했다는 말의 의미

→ 해당 사건(부분집합) 원소 중에 정말로 선택된 표본이 있다는 사실을 알게 되었다는 의미

→ 해당 사건이 내포하는 명제/가설/주장이 진실임을 알게 되었다는 의미

→ 지금까지 모르고 있던 추가적인 정보가 들어왔음을 의미

(Bayesian PoV example)

Say : 불투명 컵 안에 주사위를 넣고 굴렸다고 가정.

- "주사위의 눈금이 짝수가 나오는 사건이 발생했다"

= 컵을 들어서 주사위 눈금을 확인해보니 주사위의 눈금이 짝수였다는 사실을 알게된 것

"검진 결과 암에 걸렸을 확률은 90%입니다."

frequentist
doctor

Bayesian
patient

"이런 증상을 가진 환자는
경험상, 통계상
9할이 암에 걸린 환자였다"

"저 의사가 주장하는
내가 암에 걸렸다는 가설의
신뢰도는 90%이다"