Week 10. 순열과 조합, 확률, 확률변수, 확률분포, 베이지안
인공지능에 필요한 기초통계 지식에 대하여 학습한다.

10.1 순열과 조합
경우의 수를 세는 방법
① 순열 (permutation) ← 순서를 고려하여 나열하는 경우의 수
ex) 1 2 3 4 5 카드 5장, 3장 택하여 나열 → $5 \times 4 \times 3 = 60$

서로 다른 n개에서 k개를 택하여 순서대로 나열한 순열의 수 $_nP_k$

$$_nP_k = n(n-1)\cdots(n-k+1) \quad (k \le n)$$

* 특히 k=n일 때, $_nP_n = n(n-1)\cdots 2\cdot 1 = n!$ (n factorial)

$$_nP_k = n(n-1)\cdots(n-k+1) = \frac{n(n-1)\cdots 2\cdot 1}{(n-k)\cdots 2\cdot 1} = \frac{n!}{(n-k)!}$$

② 조합 (combination) — 순서와 상관없이 선택하는 경우의 수
ex) 1 2 3 4 5 카드 5장, 순서없이 3장 뽑기 → $5 \times 4 \times 3 / 3! = 10$
* 3!으로 나누는 것의 의미 = 뽑은 3장은 모두 "같은순서"로 보겠다는 의미

$$_nC_k = \frac{_nP_k}{k!} = \frac{n!}{(n-k)!\,k!} = \binom{n}{k} \quad (k \le n)$$

● 중복을 허락한다면?
(중복순열) 서로 다른 n개에서 중복을 허락하여 k개 뽑고 순서대로 나열, $_n\Pi_k = n^k$
(중복조합) 서로 다른 n개에서 중복을 허락하여 k개를 택하는 방법의 수 $_nH_k = {}_{n+k-1}C_k$

10.2 확률

S(sample space, 표본공간): 동전던지기 {앞면, 뒷면}, 정육면체주사위 {1, 2, 3, 4, 5, 6}

S에 속하는 특정사건의 집합을 A라 할 때, $P(A) = \frac{n(A)}{n(S)}$ ··· A 갯수 / ··· S 갯수

① 수학적 확률 $P(A) = \frac{\text{사건 A 경우의 수}}{\text{표본공간 S 경우의 수}} = \frac{n(A)}{n(S)}$

② 기하학적확률 $A \subset S$인 영역 A에 속하는 확률

$$P(A) = \frac{\text{영역 A 크기}}{\text{영역 S 크기}}$$

③ 통계적 확률과 대수의 법칙 (Law of large number)

n번의 시행동안, 특정 사건 A가 일어난 횟수가 k번이면, A의 통계적 확률을 $\frac{k}{n}$ 라 할수있다.
그러나 시행횟수 n이 충분히 커지면 통계적 확률은 수학적 확률과 같아진다.

$$\lim_{n\to\infty} \frac{k}{n} = P(A) = \frac{n(A)}{n(S)}$$

확률은 다음성질을 만족한다. 사건 A의 확률을 P(A)라 하면

1. 표본공간 S에서 임의의 사건 A에 대하여 $0 \le P(A) \le 1$이 성립한다.
2. 표본공간 S에 대하여 $P(S) = 1$ (표본공간 전체의 확률은 1)이 성립한다.
3. 공사건 ϕ에 대하여 $P(\phi) = 0$이 성립한다.
4. 두 사건 A, B가 동시에 발생하지 않는 배반사건이면 다음이 성립한다.

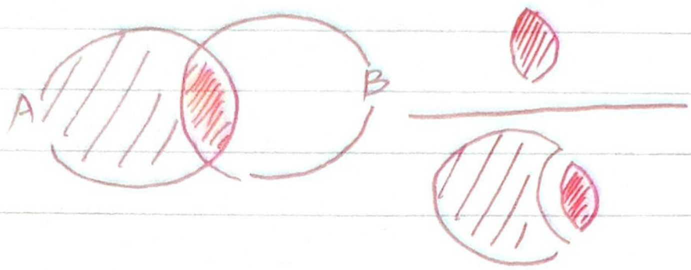
$$P(A \cup B) = P(A) + P(B)$$

5. 사건 A가 일어나지 않는 경우를 A^c라 하면 $P(A^c) = 1 - P(A)$

10.3 조건부 확률

사건 A가 일어났다는 조건 하에서 사건 B가 일어날 확률
= 사건 A에 대한 사건 B의 조건부확률 = $P(B|A)$

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



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

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

조건부 확률의 정의로부터 다음의 곱셈정리를 얻을 수 있다.

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

또한, 일반적으로 사건 $A_1, A_2, \cdots A_n$에 대하여 다음이 성립한다.

$$P(A_1)P(A_2|A_1)P(A_3|A_1\cap A_2)\cdots P(A_n|A_1\cap A_2\cdots\cap A_{n-1})$$

$$= \frac{\cancel{P(A_1)}}{1}\cdot\frac{\cancel{P(A_1\cap A_2)}}{\cancel{P(A_1)}}\cdot\frac{\cancel{P(A_1\cap A_2\cap A_3)}}{\cancel{P(A_1\cap A_2)}}\cdots\frac{P(A_1\cap A_2\cdots\cap A_n)}{\cancel{P(A_1\cap A_2\cdots A_{n-1})}}$$

$$= P(A_1\cap A_2\cdots\cap A_n)$$

10.4 베이즈 정리

(베이즈 정리(Bayes' theorem))

- 주어진 조건에서 어떠한 현상이 실제로 나타날 확률을 구하는 방법
- 불확실성 하에서 의사결정 문제를 수학적으로 다룰 때 중요하게 이용

"정보와 같이 눈에 보이지 않는 무형 자산이 지닌 가치를 계산할 때 유용"

용어

1. 사전확률(prior probability)

   관측자가 이미 알고있는 사건으로부터 나온 확률, $P(A)$ = A에 대한 사전확률

2. 사후확률(posterior probability)

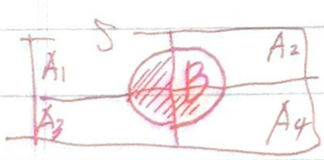
   사전확률과 대비되는 개념, 실제의 데이터나 조건이 부과되었을 때 기대되는 조건부 확률

   - 이미 발생한 사건의 '원인'이 불확실한 상황을 도식화, $P(A|B)$
   - B는 이미 일어난 사건 // B를 관측한 후 원인사건 A를 확률로 따졌다는 의미

⇒ 베이즈정리는
사전확률과 사건으로부터 얻은 자료를 사용하여 사후확률을 추출해내는 것

$\{A_1, A_2, \cdots A_n\}$이 표본공간 S의 분할(partition)을 이룬다고 하자. 그러면 임의사건 B에 대하여,

$$B = S \cap B = (A_1 \cup A_2 \cup \cdots A_n) \cap B$$
$$= (A_1 \cap B) \cup (A_2 \cap B) \cup \cdots \cup (A_n \cap B)$$



이 때 $A_i \cap B$ $(i = 1, 2, \cdots n)$는 서로 배반(exclusive)이다. 따라서

$$P(B) = P(A_1 \cap B) + P(A_2 \cap B) + \cdots + P(A_n \cap B)$$

한편, 확률의 곱셈정리로부터 아래 전확률 공식(Law of Total Probability)을 얻는다.

$$P(B) = P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + \cdots + P(A_n)P(B|A_n)$$

또한, 임의의 i에 대한 조건부확률 $P(A_i|B) = \frac{P(A_i \cap B)}{P(B)}$ 에

$P(A_i \cap B) = P(A_i)P(B|A_i)$ 와 전확률공식을 대입하면 다음식을 얻으며,

이를 베이즈정리 (Bayes' theorem) 라고 한다.

$$\boxed{P(A_i|B)} = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i) \cdot P(B|A_i)}{P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + \cdots + P(A_n)P(B|A_n)}$$

사건 $P(A_i)$의 사후확률 (after we get 'B' information/clue)

예제 3대의 기계 A, B, C 가 공장 생산량 50%, 30%, 20%를 각각 차지.

이들 기계가 불량품을 생산할 비율은 각각 4%, 3%, 2%

Q1. 한 제품을 임의로 선택할 때, 그 제품이 불량품일 확률

Q2. 불량품이 기계 C에 의하여 생산될 확률

Q1.

사건의 정의

C = 구입한 1개의 제품이 기계 C로 생산된 제품인 사건

X = 기계 C로 생산된 제품이 불량품인 사건

(제품을 생산하는 사건) $A \cup B \cup C = S$ 이고, $P(A) = .5$, $P(B) = .3$, $P(C) = .2$

(불량품을 생산하는 사건) $X = (A \cap X) \cup (B \cap X) \cup (C \cap X)$

(불량품을 생산하는 확률) $P(X|A) = .04$, $P(X|B) = .03$, $P(X|C) = .02$

전확률 공식에 의해,

$$P(X) = P(A)P(X|A) + P(B)P(X|B) + P(C)P(X|C)$$
$$= 0.5 \times 0.04 + 0.3 \times 0.03 + 0.2 \times 0.02 = 0.033$$

베이즈 정리에 의해, 불량품 중 기계 C가 생산한 확률이 불량품일 확률

$$P(C|X) = \frac{P(C)P(X|C)}{P(A)P(X|A) + P(B)P(X|B) + P(C)P(X|C)} = \frac{0.2 \times 0.02}{0.5 \times 0.04 + 0.3 \times 0.03 + 0.2 \times 0.02}$$

$$= \boxed{\frac{4}{33}} \approx \boxed{12\%}$$

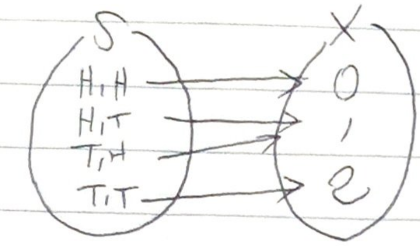
10.5 확률변수

"동전 2개 던지기"

(H, H) (H, T) (T, H) (T, T)

$\frac{1}{4}$ $\frac{1}{4}$ $\frac{1}{4}$ $\frac{1}{4}$

뒷면이 나오는 동전의 개수 = X 라 하면,



? 확률변수 random variable

컴퓨터 프로그래밍에서의 변수와 같은 개념, "어떤 값을 취하느냐가 확률적으로 결정되는 변수"

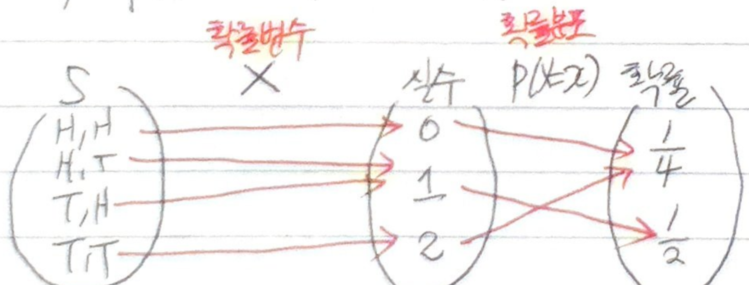
⇒ 표본공간의 모든 표본에 대해 어떤 실수값을 대응(할당)시킨 것

* 확률변수는 영문자 대문자로 쓰고 그 변수가 취할 수 있는 값 하나하나에 대해 소문자로 쓴다.

10.6 이산확률분포

확률변수 X가 연속적이지 않은 값 $x_1, x_2, \cdots x_n$을 취할 때, X를 이산확률변수라 하고, 각각의 x_i에 대하여 $X=x_i$일 확률 $P(X=x_i)$을 할당한 것을 이산확률분포라 한다.

X	x_1	x_2	$\cdots$	x_n	sum
P	$P(X=x_1)$	$P(X=x_2)$	$\cdots$	$P(X=x_n)$	1



이 때, 이산확률변수 X가 $x_1, x_2, \cdots x_n$의 값을 취할 때 확률 $P(X=x_i)$를 대응시키는 함수 $f(x)$를 확률변수 X의 확률질량함수(PMF, Probability Mass Function)라 한다.

$$f(x) = \begin{cases} P(X=x_i), & x = x_1, x_2, \cdots x_n \\ 0, & \text{그 외의 값} \end{cases}$$

(확률질량함수의 성질)

① $0 \le f(x_i) \le 1 \quad (i=1,2,\cdots n)$

② $f(x_1)+f(x_2)+\cdots+f(x_n) = \sum_{i=1}^{n} f(x_i) = 1$

③ X가 이산확률변수일 때 $a \le X \le b$일 확률

$$P(a \le X \le b) = \sum_{a \le x \le b} f(x)$$

10.7 연속확률분포

* 확률변수 X가 어떤 범위에 속하는 모든 실수를 취할 때, X를 연속확률변수라 한다.
* 연속확률변수의 경우, 특정한 실수 x 값을 취할 확률 $P(X=x)$ 가 항상 0이므로, 확률질량함수를 이용하여 확률분포를 나타내는 것은 의미가 없다.

→ 그래서 확률밀도함수 (PDF, Probability Density Function)를 새로 정의하여 연속확률분포를 나타낸다.

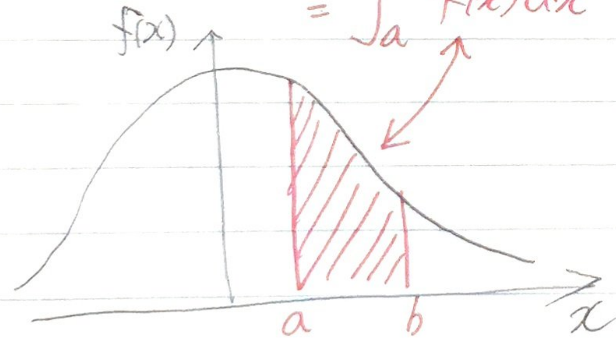
연속확률변수 X에 대하여 함수 $f(x)$가 다음 성질을 만족하면 $f(x)$를 확률밀도함수라 한다.
(확률질량함수 성질에 대응되는 개념)

① 모든 실수 x에 대하여 $f(x) \geq 0$ 이다.

② $\int_{-\infty}^{\infty} f(x)\,dx = 1$

③ $P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b)$

$= \int_a^b f(x)\,dx$



? '밀도(density)'라는 단어가 왜 쓰였을까?

- 확률을 일종의 양 (질량)으로 보고, 구간의 길이를 일종의 부피로 본다면,
(확률 / 구간의 길이) × (구간의 길이) = (확률) 이므로,
(확률/ 구간의 길이)는 (질량/부피)가 되어 '밀도'를 의미하게 된다.
질량을 부피로 나눈 것 — 얼마나 빽빽하게 차 있는가?