[K-MOOC] 인공지능에 필요한 기초수학 입문(High school)
(Introductory Math for Artificial Intelligence)

이상구 | 성균관대학교

10주차. 순열과 조합, 확률, 확률변수, 확률분포, 베이지안

* 순열(permutation) : 순서를 고려하여 나열하는 경우의 수를 의미한다. 서로 다른 n개에서 k개를 택하여 순서대로 나열한 순열의 수를 ${}_nP_k$ 로 쓰고 다음 공식에 의해 계산한다.

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

* 조합(combination) : 순서와 상관없이 선택하는 경우의 수를 말한다. 서로 다른 n개에서 k개를 택하는 조합의 수를 ${}_nC_k$ 와 같이 나타내고 다음 공식에 의해 계산한다.

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

* 중복순열(repeated permutation) : 서로 다른 n개에서 중복을 허용하여 k개를 택하여 순서대로 나열한 경우의 수이고 다음 공식에 의해 계산한다.

$${}_n\Pi_k = n^k$$

* 중복조합 (repeated combination) : 서로 다른 n개에서 중복을 허용하여 (순서없이) k개를 택하는 경우의 수이고 다음 공식과 같이 계산할 수 있다.

$${}_nH_k = {}_{n+k-1}C_k$$

* 확률(probability) : 어떤 실험이나 관찰에서 각 경우가 일어날 가능성이 같다고 할 때, 일어나는 모든 경우의 수를 $n(S)$, 사건 A가 일어나는 경우의 수를 $n(A)$라고 하면 사건 A가 일어날 확률 $P(A)$는 다음과 같다.

$$P(A) = \frac{\text{사건 } A\text{가 일어나는 경우의 수}}{\text{일어날 수 있는 모든 경우의 수}} = \frac{n(A)}{n(S)}$$

* 조건부 확률 : 어떤 사건 A가 일어났다는 조건하에서 사건 B가 일어날 확률을 사건 A에 대한 사건 B의 조건부확률(conditional probability)이라 하고 $P(B|A)$로 표시하며 다음과 같이 정의한다.

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

* 베이즈 정리(Bayes' theorem)는 주어진 조건에서 어떠한 현상이 실제로 나타날 확률을 구하는 방법이다.

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

베이즈 정리에서 $P(A_i)$를 사건 A_i의 사전확률, $P(A_i|B)$를 사건 A_i의 사후확률이라 한다.

* 확률변수(random variable) : 표본 공간의 모든 표본에 대해 어떤 실수 값을 대응시킨(할당한) 것

* 이산확률변수, 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.

* 확률질량함수 : 이산확률변수 X가 $x_1, x_2, \ldots, x_n$의 값을 취할 때 확률 $P(X=x_i)$을 대응시키는 함수 $f(x)$를 확률변수 X의 확률질량함수(probability mass function)라 한다.

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

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

* 확률밀도함수(probability density function) : 연속확률변수 X에 대하여 함수 $f(x)$가 다음 성질을 만족하면 $f(x)$를 X의 확률밀도함수라 한다.
  ① 모든 실수 x에 대하여 $f(x)\geq 0$ 이다.
  ② $\int_{-\infty}^{\infty} f(x)\,dx=1$
  ③ $P(a\leq X\leq b)=P(a<X\leq b)=P(a\leq X<b)=P(a<X<b)=\int_a^b f(x)\,dx$