

# [7주차] 확률변수와 확률분포

## Random Variable is a real-valued function

데이터를 이해하기 위한 확률변수의 이해

$\omega$	A	B	C	D
1	$a_{11}$	$b_{21}$	$c_{31}$	
2	$a_{12}$	$b_{22}$	$c_{32}$	
3	$a_{13}$	$b_{23}$	$c_{33}$	
4	$a_{14}$	$b_{24}$	$c_{34}$	
$\vdots$	$\vdots$	$\vdots$	$\vdots$	

⇒ 확률 표본(Random sample)

= Data

Random Variable

A random variable is a real valued function defined on probability space;

$(\Omega, \mathcal{F}, p)$

where,

$\Omega$  : sample space

$\mathcal{F}$  : set of events

$p$  : probability

$X : \Omega \rightarrow \mathbb{R}$

Ex) 동전 1개 던지는 실험

$\Omega = \{H, T\}$  ; 여기서  $\{H\}$ ,  $\{T\}$ 를 표본공간의요소

$\mathcal{F} = \{\{H\}, \{T\}, \phi, \Omega\}$  : 사건들의 집합

$p : \mathcal{F} \rightarrow [0, 1]$

$X = 0$ , IF  $\{T\}$

1, IF  $\{H\}$

사례 = 동전 2개를 동시에 던지는 실험에서  
 나타낼 변수  $X$ 를 어떻게 나타낼 수 있을까?

① 표현할 변수  $X$ 를 정의하자.

$$X_i = \begin{cases} 0, & \text{if } \{T\} \\ 1, & \text{if } \{H\} \end{cases}, \quad i=1, 2$$

$$X = \sum_{i=1}^2 X_i = X_1 + X_2 \text{로 정의} *$$

② 표현할 변수의 정의역을 고려하자 !!

표현할 변수가 나타내는 함수로 정의되기 때문에,  
 정의구역을 명확히 생각해야 함

$$\Omega = \{HH, HT, TH, TT\}$$

③ 표현할 변수는 어떤 실수값을 갖게 되는가?

**확률밀도함수:** 확률변수  $X$ 가 실수값  $x$ 를 취할때의 확률값

$$X: \Omega \rightarrow \mathbb{R},$$

where  $X_i = \begin{cases} 0, & \{T\} \\ 1, & \{H\} \end{cases}, \quad i=1, 2$

$$X(\{HH\}) = X_1(\{H\}) + X_2(\{H\}) = 1 + 1 = 2$$

$$X(\{HT\}) = X_1(\{H\}) + X_2(\{T\}) = 1 + 0 = 1$$

$$X(\{TH\}) = X_1(\{T\}) + X_2(\{H\}) = 0 + 1 = 1$$

$$X(\{TT\}) = X_1(\{T\}) + X_2(\{T\}) = 0 + 0 = 0$$

$$\Rightarrow \therefore X(\cdot) = x = 0, 1, 2,$$

$X$ 의 확률 분포 (또는 확률 밀도 함수)

$X \sim f(x)$  라는 확률 밀도 함수를 갖는다면,

③ 표현할 변수  $X$ 가 실수값  $x$ 를 취할때의 확률값

$$\textcircled{1} X = x_1, x_2, x_3, \text{ then } f(x_1), f(x_2), f(x_3)$$

$$\text{Such that, } f(x_1) = P(X=x_1)$$

$$f(x_2) = P(X=x_2)$$

$$f(x_3) = P(X=x_3)$$

그럼, 앞서 소개한 동전 2개를 던지는 실험에서,  
 앞면의 개수를 표현할 변수  $X$ 라 정의한다면,  
 이때의  $X$ 의 확률 분포는??

Review : 동전 2개,  $X_i = \begin{cases} 0, & \{T\} \\ 1, & \{H\} \end{cases}, i=1,2$

⇒ 확률변수  $X = X_1 + X_2 \Rightarrow X: \Omega \rightarrow \mathbb{R}$

⇒  $X$ 의 표본공간  $\Omega = \{\{HH\}, \{HT\}, \{TH\}, \{TT\}\}$

⇒  $X$ 의 치역역사, 즉 가능한 실수값  $= \{0, 1, 2\}$

즉,  $X=0, 1, 2$ .

⇒  $X$ 의 확률분포는?



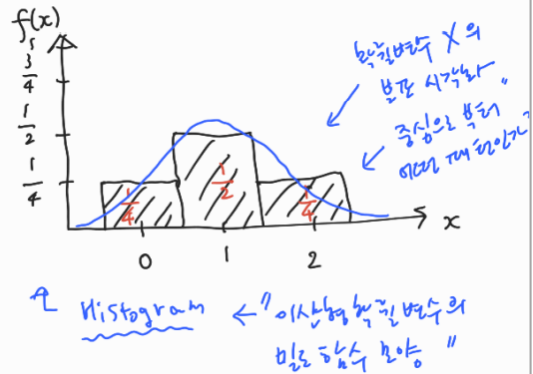
$$\begin{aligned} f(0) &= P(X=0) = P(\{\omega \in \Omega; X(\omega)=0\}) \\ &= P(\{TT\}) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} \end{aligned}$$

$$\begin{aligned} f(1) &= P(X=1) = P(\{HT\} \cup \{TH\}) \\ &= P(\{HT\}) + P(\{TH\}) \\ &= \frac{1}{4} + \frac{1}{4} = \frac{1}{2} \end{aligned}$$

$$f(2) = P(X=2) = P(\{HH\}) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

$X$ 의 확률분포표와 시각화

$X$	0	1	2	계
$P(X=x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	1



확률변수는 정략적인 실수값을 가지며 사칙연산이 가능해야합니다.

확률변수의 중심위치는 **기댓값**

기댓값으로부터 퍼져있는정도 **분산**