

# Random Variable? 그게 뭔가요? (확률 변수???)

소프트웨어 공대 강의

노기섭 교수

(kafa46@cju.ac.kr)

# 가장 헷갈리는 개념 Random Variable $X$

딥러닝 수학 시리즈를 시작하게 된 계기?

딥러닝?

머신러닝?

확률 분포?

시계열?

확률변수?



이미지 출처: <https://smarticons.co/blog/understanding-the-meaning-of-question-mark-icons/>

# Random Variable

## ■ Constant vs. variable

- 우리가 알고 있는 상수(constant)?

$$\pi = 3.141 \dots$$

$$e = 2.718 \dots$$

- 우리가 알고 있는 변수(variable)?

$$3x + 2y = 5$$

$$y = 2x + 1$$

- variables:  $x, y$

## ■ Random variable (확률변수)

- 확률에 따라서 변하는 수?
  - E.g., Coin toss
- 어떤 사건에 대하여 편리한 표현은 없을까?

## 임의의 실험

실험 종류	가능한 결과	실제 결과 ( $X$ )
동전 2개 던지기	(앞, 앞), (앞, 뒤) (뒤, 앞), (뒤, 뒤)	?
몸무게 측정	$(0, \infty]$	?
2개 주사위를 던졌을 때 합	$\{2, 3, \dots, 12\}$	?
$\vdots$	$\vdots$	$\vdots$

실제 결과  $X$  는 실험한 이후에 알 수 있음... $\pi$

실제 결과  $X$  는 확률에 따라 표시할 수 있을 것

$$\left\{ \begin{array}{l} P(\text{동전 2개 던지기를 했을 때 결과}) \\ P(\text{몸무게를 측정했을 때 값}) \\ P(\text{주사위 2개를 던졌을 때 더한 값}) \\ \vdots \end{array} \right.$$

# Conceptualizing

## ■ Random variable (확률변수)



$P(\text{동전 2개를 던졌을 때 결과 첫번째 동전은 앞면, 두번째 동전은 뒷면})$

너무 길게 써야 하는데? 귀찮아!  
그냥  $X$  라고 쓰자!!!

$P(X) \rightarrow$  어라! 가능한  $X$  값들을 어떻게 지정하지?

그냥 함수로 만들자!

$X: \Omega \rightarrow E$

$\Omega$ : sample space

$E$ : measurable space

$P(X = 1)$

둘 다 가능?

$\Omega$	$X$	$E$	$E$
앞, 앞	$\rightarrow$	0	$\pi e^{3 \times 0}$
앞, 뒤	$\rightarrow$	1	$\pi e^{3 \times 1}$
뒤, 앞	$\rightarrow$	2	$\pi e^{3 \times 2}$
뒤, 뒤	$\rightarrow$	3	$\pi e^{3 \times 3}$

# Summary & More Example

## ■ Random variable 정리하면

- 어떤 실험 결과가 확률적으로 나타날 때
- 그리고 실험 결과로 나올 수 있는 모든 경우의 수(sample space,  $\Omega$ )를 알고 있을 때
- 실험 결과를 특정 값으로 짝지어 주는 함수 (mapping function 또는 measurable function)
- **Random variable** 은 함수다!

일반적으로 random variable은  
알파벳 대문자로 표시함 ( $X, Y, Z, \dots$ )

## ■ 또 다른 예시: 주사위 2개를 던지는 실험



$\Omega$	$X$	E
1, 1	→	2
1, 2	→	3
⋮		⋮
2, 1	→	3
⋮	→	⋮
6, 6		12

함수 정의:  $X$ (주사위 2개 던지기)=두 값의 합

$P$ (주사위 2개를 던졌을 때 두 값의 합이 3)

$$P(X = 3) = \frac{\text{집합 } E \text{ 에서 값이 } 3 \text{ 인 개수}}{\text{집합 } E \text{ 의 원소 개수}},$$

$P_X(3)$ 으로 표현하기도 하고,

축약해서  $P(3)$ 으로 표시할 수도 있음

# Random Variable과 확률

## ■ 확률(Probability)?

- 확률(確率, probability)은 어떤 일이 일어날 가능성이다. 비율이나 빈도로 측량해 나타낼 수 있다.  
([http://www.ktword.co.kr/test/view/view.php?m\\_temp1=1644](http://www.ktword.co.kr/test/view/view.php?m_temp1=1644), 정보통신기술용어해설집)

## ■ Random variable 출력

- $X: \Omega \rightarrow E, \text{where } \Omega: \text{sample space \& } E: \text{measurable space}$
- 일반적으로  $E$ 는 실수 값

## ■ Random variable과 확률의 연관성

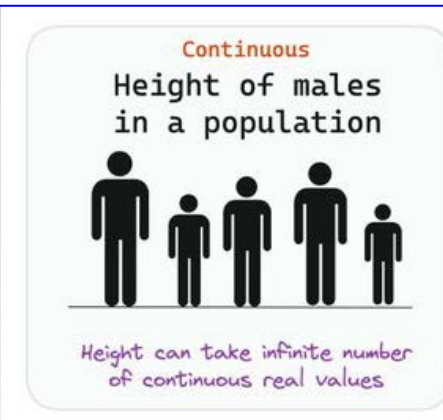
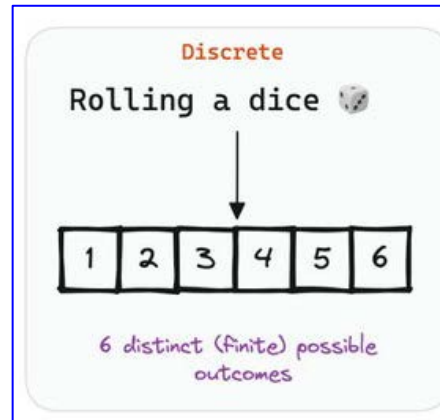
- Random variable에 해당하는 값이 얼마나 자주 발생하는지 관찰  
(예) Random variable  $X$ 의 출력(실수)값  $x$  얼마나 자주 발생하는가?  $P(X = x)$
- 실수 값  $x$ 는 확률 값을 알아내기 위한 변수(변하는 값)으로 해석할 수도 있을 것
- Random variable에 "변수(variable)" 라는 이름이 붙은 이유

# Types of Random Variables

## ■ Random variable 종류?

### - Discrete

- 실험 결과를 이산적인 값으로 맵핑



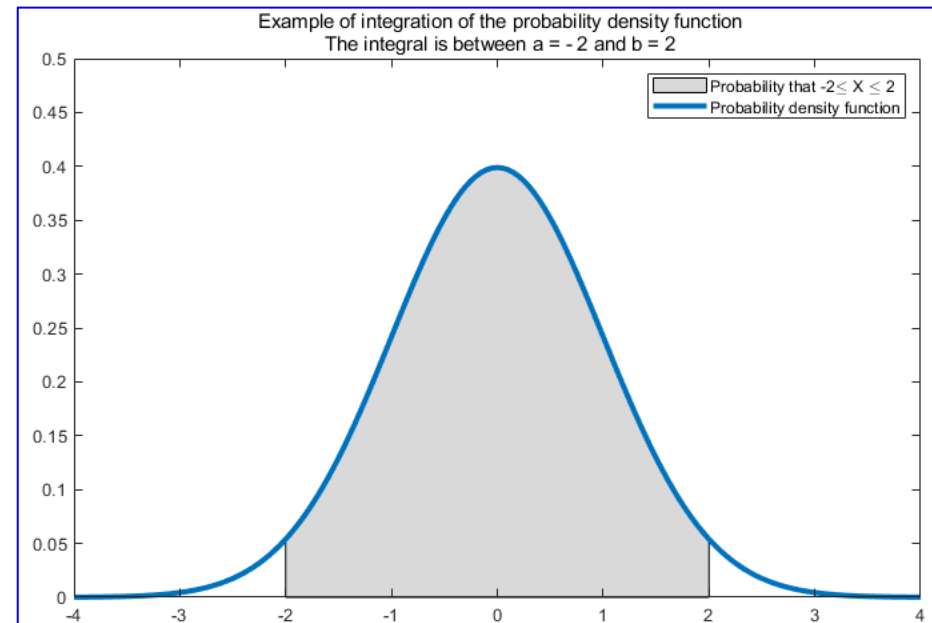
이미지 출처:  
[https://twitter.com/akshay\\_pachaa/status/1668234250343305216](https://twitter.com/akshay_pachaa/status/1668234250343305216)

### - Continuous

- 실험 결과를 특정 구간의 모든 실수 값을 취하는 연속 구간의 값으로 맵핑

$$P(X \in [a, b]) = \int_a^b f_X(x) dx,$$

where  $f_X(x)$  is probability density function



이미지 출처: <https://faculty.nps.edu/rbassett/book/continuous-random-variables.html>

# More Example: Continuous Random Variable

## ■ Discrete random variable

- 정확한 값을 알고 있을 경우 (**distinct or separated values**)

$$X = \begin{cases} 1 & \text{Head} \\ 0 & \text{Tail} \end{cases}, \text{where fair coin toss}$$

$X$  = 랜덤하게 1명의 학생을 선택했을 경우 출생년도

$X$  = 우리나라에서 내일 태어날 개미 숫자

## ■ Continuous random variable? → **항상 헛갈리는 개념**

- 정확한 값은 있지만, 정확히 모를 경우 (**any value in interval**)

$Y$  = 청주에 있는 고양이의 (**정확한**) 평균 몸무게

$Y = 37.234876 \dots \text{Kg} \rightarrow$  어떤 구간에 있을 확률로

표시할 수 밖에 없음 (예: 37.2 Kg ~ 37.3 Kg 사이에 있을 확률)

$Y$  = 최근 올림픽 경기에서 100미터 달리기 우승자의 (**정확한**) 기록

$Y = 9.5812432 \dots \text{초} \rightarrow$  어떤 구간에 있을 확률로

표시할 수 밖에 없음 (예: 9.58초 ~ 9.59초 사이에 있을 확률)





수고하셨습니다 ..^^..