

선수 지식 - 통계

확률 변수와 확률 분포

확률 변수와 확률 분포 | 딥러닝의 기초가 되는 확률 개념 알아보기

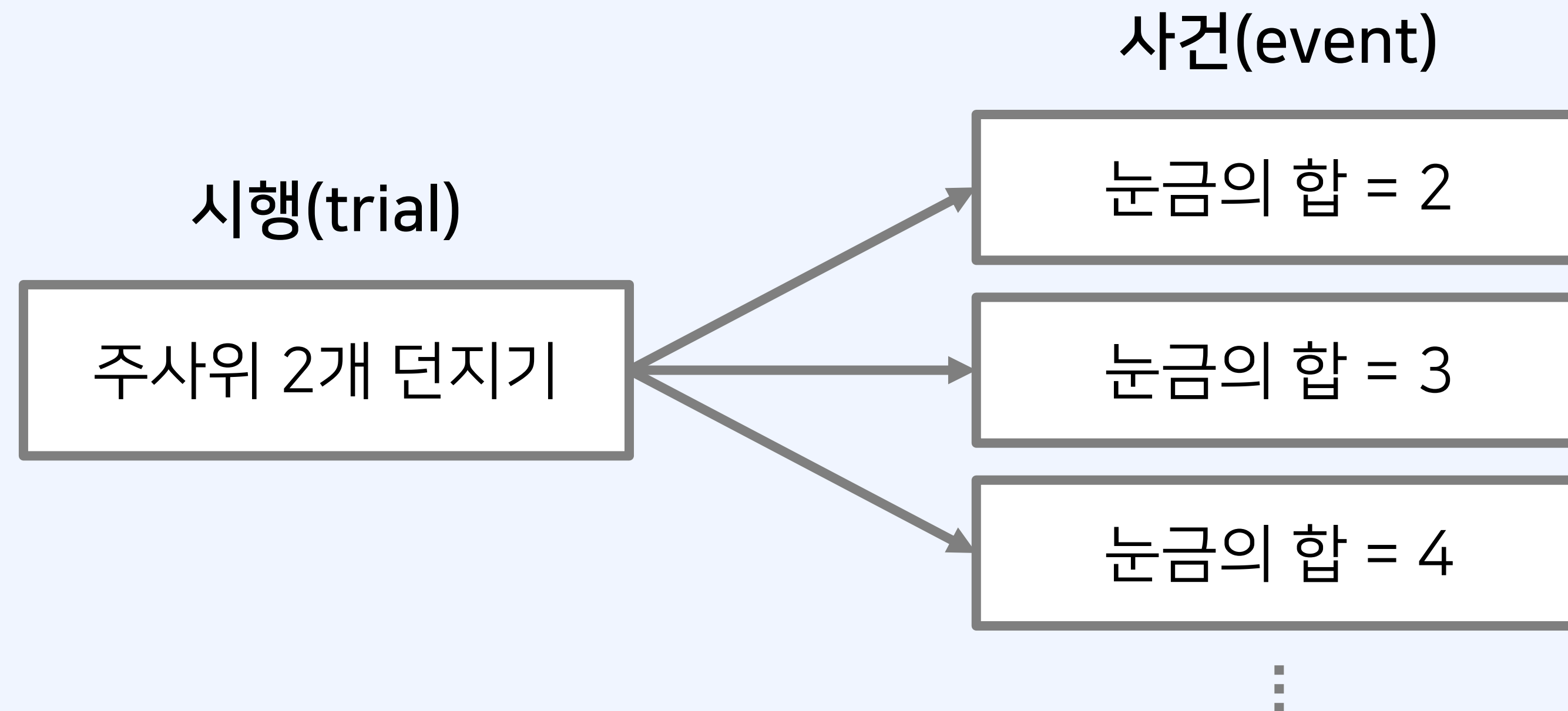
강사 나동빈

선수 지식 - 통계

확률 변수와 확률 분포

시행(Trial)과 사건(Event)

- 확률에 대하여 이해하기 위해서, 먼저 시행(trial)과 사건(event)에 대해 알아야 한다.
- 시행(trial): 반복할 수 있으며, 매번 결과가 달라질 수 있는 실험 ex) 주사위를 2개를 던지는 행동
- 사건(event): 시행에 따른 결과를 의미 ex) 눈금의 합이 7이 되는 사건



확률(Probability)

- 확률은 어떠한 사건(event)이 일어날 가능성을 수로 표현한 것이다.
- 주사위 2개를 던졌을 때, 눈금의 합이 7이 나올 확률을 구하고 싶다고 해보자.
- 이러한 문제를 수식으로 표현하려면 어떻게 해야할까?

확률 변수(Random Variable)

- 확률 변수란, **사건으로 인해** 그 값이 확률적으로 정해지는 변수를 의미한다.
- 주사위 2개를 던지는 시행을 할 때마다 눈금의 합이 변할 수 있다.
- 따라서 “**확률 변수 = 눈금의 합**”으로 표현할 수 있다.
- 확률 변수는 대문자 X 로 표기하고, 확률 변수가 취할 수 있는 값은 소문자 x 로 표현한다.

확률 함수(Probability Function)

- 앞서 확률 변수란, 시행할 때마다 변할 수 있는 값(눈금의 합)이라고 했다.
- **확률 함수**란, 확률 변수에 따라서 확률 값을 부여하는 함수를 말한다.
- 확률 함수는 일반적으로 P 라고 표현한다.
- 주사위 두 개 던지기(시행)을 했을 때 눈금의 합이 3이 나올 확률은 $2/36$ 이다.
- 눈금의 합이 3이 되는 경우로는 (1, 2)와 (2, 1)의 두 가지 경우가 존재한다.
- 이를 다음과 같이 표현할 수 있다.
- $P(X = 3) = 2/36$

확률 변수(Random Variable) 예시

- 앞면에 1, 뒷면에 0이 쓰여 있는 2개의 동전이 있다고 해보자.
- 시행: 동전 2개를 동시에 던지기
- 확률 변수 X : 두 눈금의 합
- 이때 가능한 X 의 값으로는 0, 1, 2가 있다.
- 시행할 때마다 X 값이 달라질 수 있으므로, 확률적으로 변할 수 있다고 하여 X 를 확률 변수라 한다.

확률 변수(Random Variable) 예시

- 2개의 동전을 던지는 시행에 대하여:
- 수학적으로, 각 사건에 대한 확률은 다음과 같이 표현할 수 있다.
- 눈금의 합이 0인 사건이 발생할 확률: $P(X = 0) = 1/4$
- 눈금의 합이 1인 사건이 발생할 확률: $P(X = 1) = 1/2$
- 눈금의 합이 2인 사건이 발생할 확률: $P(X = 2) = 1/4$

시행: 동전 2개를 동시에 던지기
확률 변수 X : 두 눈금의 합

딥러닝 분야에서의 사건(Event)

- 우리가 흔히 얻을 수 있는 데이터는 사건(event)로 이해할 수 있다.
- 이미지 분류 모델을 학습할 때는 다양한 이미지를 사용한다.
- 이때 내가 수집하여 가지고 있는 이미지를 사건(event)으로 이해할 수 있다.



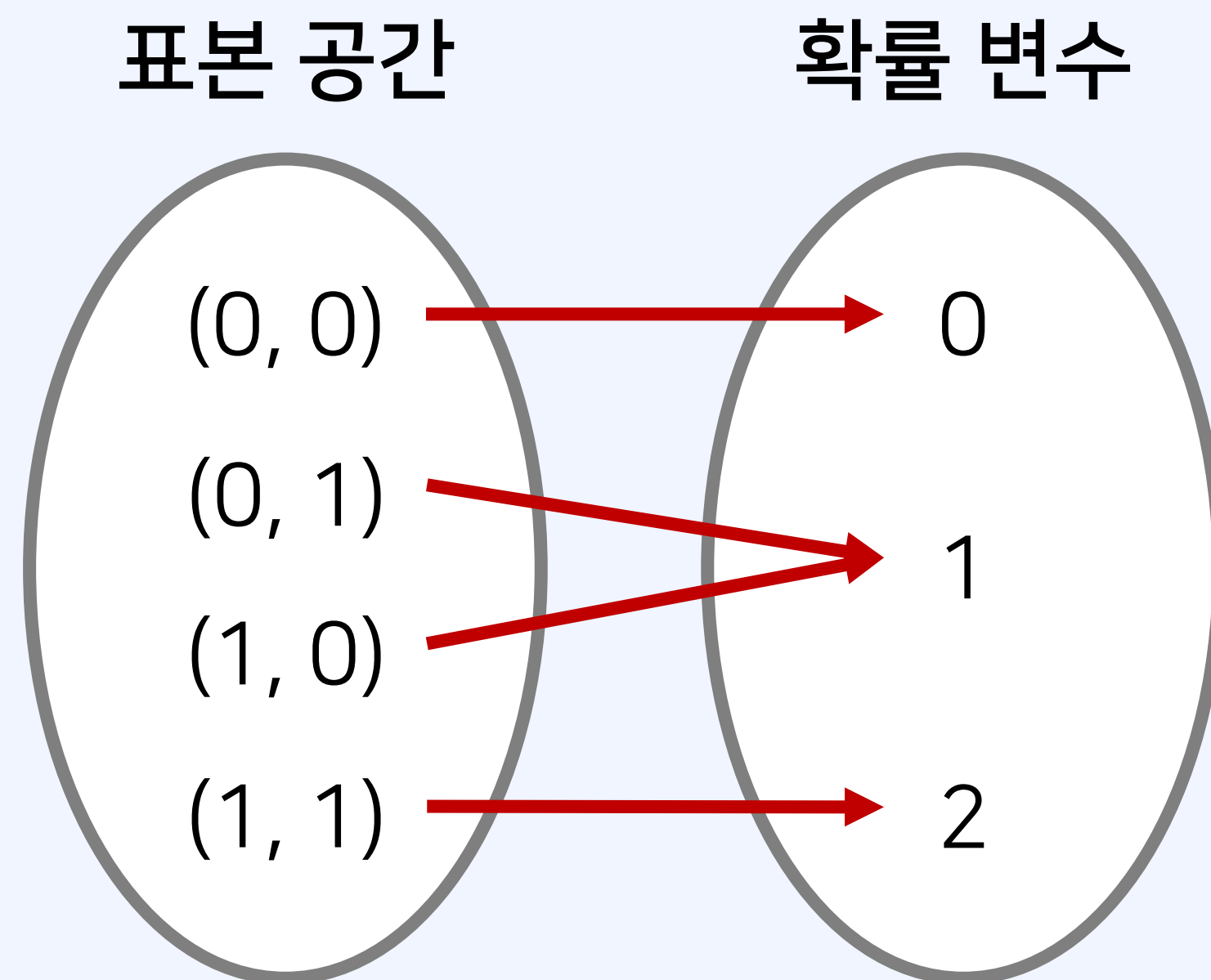
이미지 x

[생각할 거리] $P(x)$ 의 의미는?

존재하는 모든 이미지 중에서 이미지 x 가 나올 확률을 의미한다.

확률 변수는 함수다?

- 표본 공간(sample space): 발생 가능한 모든 사건들의 집합(전체 집합)
- 확률 변수는 표본 공간 안에서 특정 확률로 발생하는 사건을 특정한 수치에 대응시키는 함수다.
- 이때 확률 변수가 취할 수 있는 모든 값 x 의 집합을 상태 공간(state space)이라고 한다.



상태 공간: $\{0, 1, 2\}$

[확률 함수]

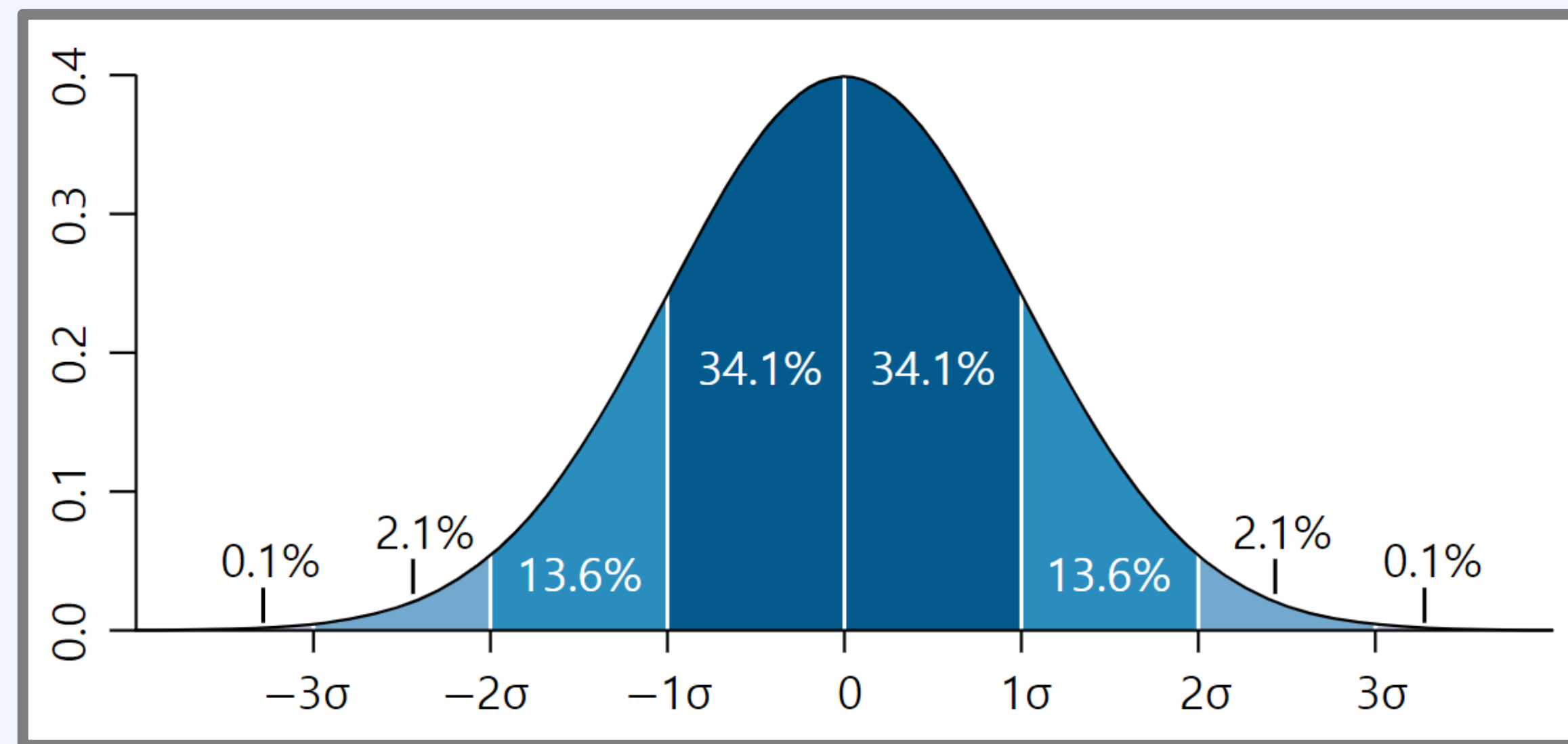
$$P(X = 0) = 1/4$$

$$P(X = 1) = 1/2$$

$$P(X = 2) = 1/4$$

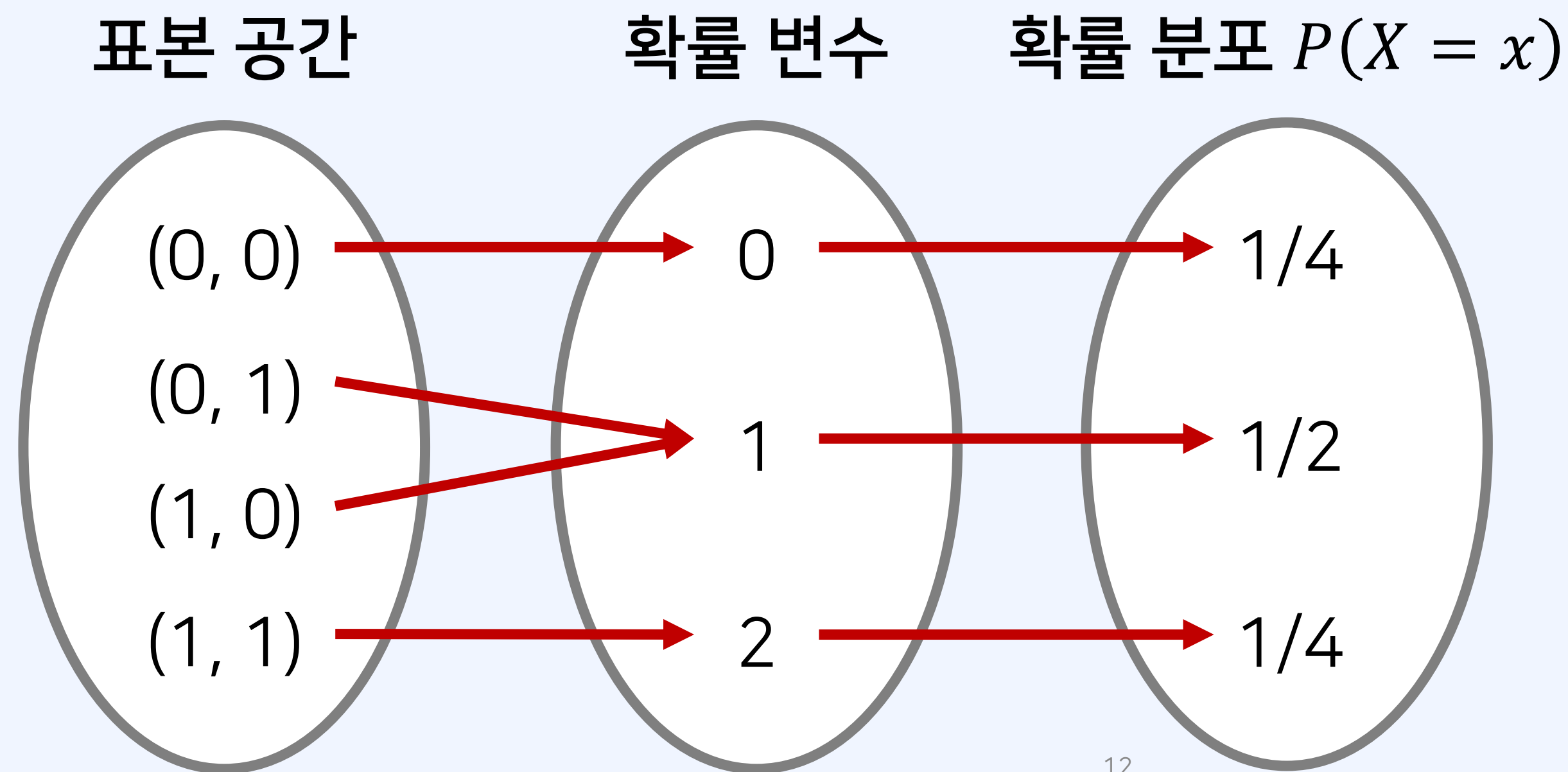
확률 분포(Probability Distribution)

- 확률 분포(probability distribution)란?
- 각 사건에 어느 정도의 확률이 할당되었는지 표현한 정보를 의미한다.
- 확률 분포를 통해 통계적인 특성을 쉽게 이해할 수 있다.



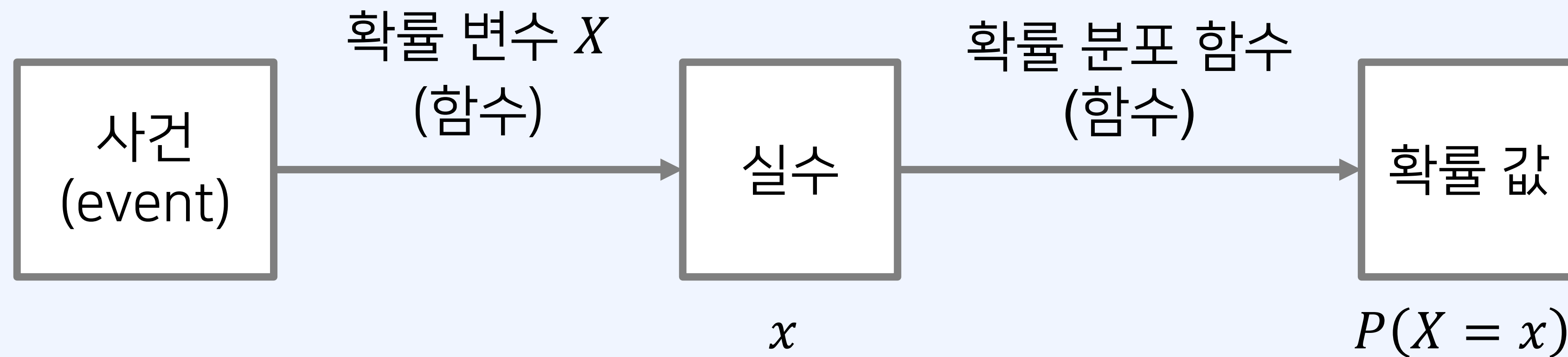
확률 분포 함수(Probability Distribution Function)

- 확률변수 X 가 가지는 값 x 에 확률 $P(X = x)$ 을 대응시키는 함수를 의미한다.
- [참고] 확률 분포 그 자체를 함수로 보고, 확률 분포 함수 P 와 같은 의미로 쓰기도 한다.
- 확률 분포 함수로 대표적인 것으로는 ① 확률질량함수, ② 확률밀도함수가 있다.
- 모든 사건에 대하여 확률 분포 함수의 값을 표현한 것을 "확률 분포"로 이해할 수 있다.



함수로 이해하기

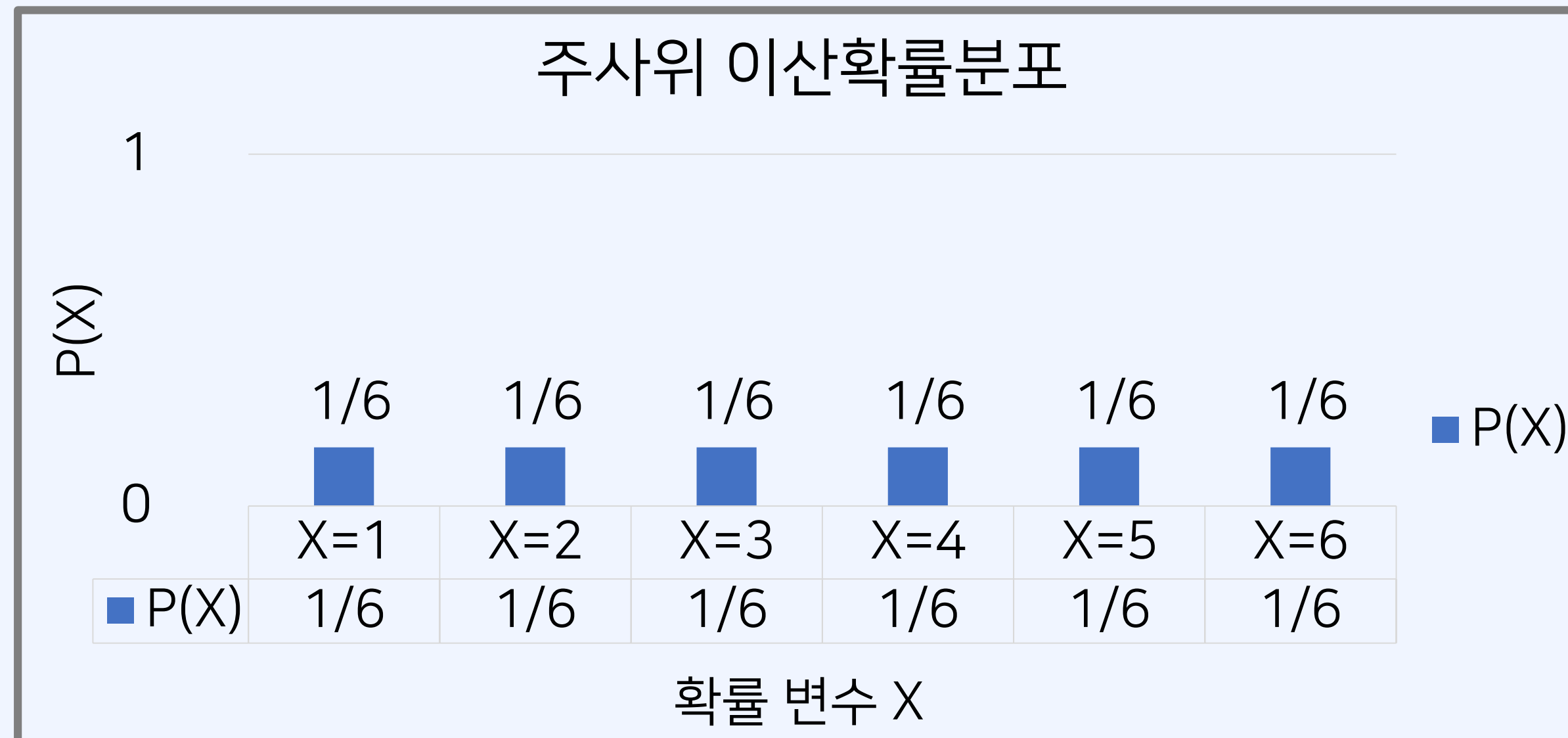
- 확률 변수와 확률 분포 함수를 모두 함수로 이해할 수 있다.



이산확률분포(Discrete Probability Distribution)

- 확률변수 X 가 취할 수 있는 모든 값을 셀 수 있는 경우, 이를 이산확률변수라고 한다.
- 이때 이산확률분포는 이산확률변수의 확률 분포를 의미한다.
- 주사위를 던졌을 때 나올 수 있는 눈금은 $\{1, 2, 3, 4, 5, 6\}$ 중 하나이다.
- 따라서, 이 경우 “주사위 눈금”의 값은 6개만 존재하므로, 이산확률변수이다.

- 주사위를 던졌을 때 나올 수 있는 눈금(수)을 확률변수 X 라고 하자.
- 확률변수 X 는 $\{1, 2, 3, 4, 5, 6\}$ 중 하나의 값을 가질 수 있다.
- $P(X = 1) = P(X = 2) = P(X = 3) = P(X = 4) = P(X = 5) = P(X = 6) = 1/6$
- 이때, 주사위 눈금 X 에 대한 확률 분포는 다음과 같다.



이산 균등 분포: 모든 확률 변수가 나올 확률이 같은 확률 분포다.

확률질량함수(Probability Mass Function, PMF)

- 확률질량함수는 이산확률변수가 특정한 값을 가질 **확률을 출력하는 함수**다.
- 확률질량함수는 이산확률분포를 표현하기 위해 사용하는 확률분포함수로 이해할 수 있다.
- 동전 2개를 동시에 던지는 시행에서 두 눈금의 합을 X 라고 하자.
- 이때, X 는 이산확률변수로, 확률질량함수 $f(x)$ 는 다음과 같이 정의할 수 있다.

$$f(0) = P(X = 0) = 1/4$$

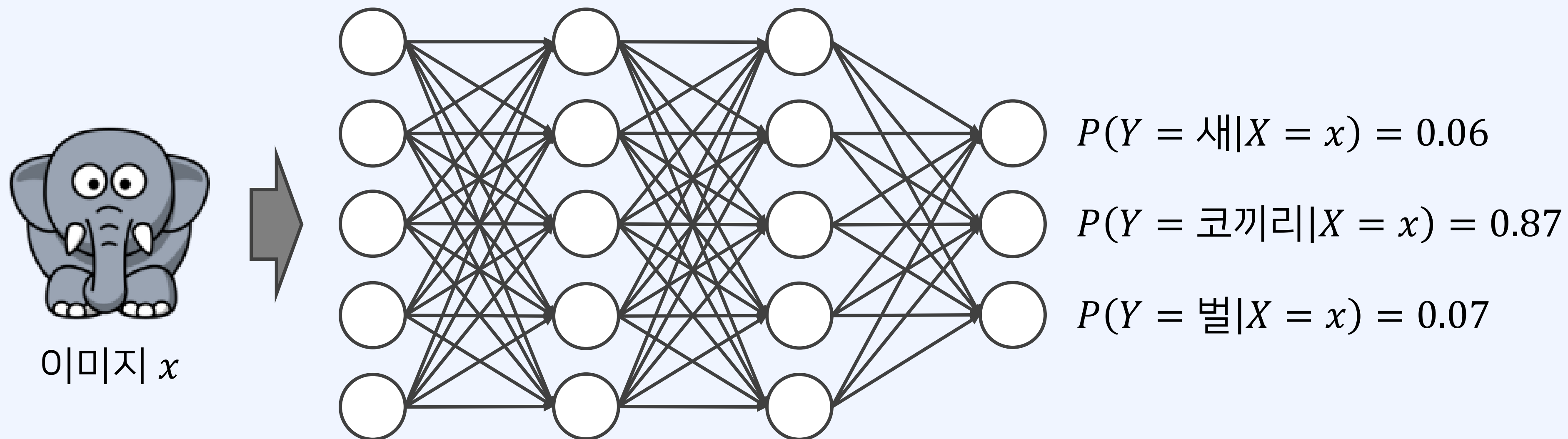
$$f(1) = P(X = 1) = 1/2$$

$$f(2) = P(X = 2) = 1/4$$

- 확률 변수 X 에 대한 확률질량함수라는 의미로 $f_X(x)$ 라고 표기하기도 한다.

확률질량함수 예시 1)

- 딥러닝을 공부할 때, 확률 분포에 대해서 꼭 이해할 필요가 있다.
- 분류(classification) 모델의 출력은 확률 분포에 해당한다.
- 이미지 x 가 주어졌을 때 클래스 y 의 확률을 $P(y|x)$ 로 표현할 수 있다.
- 확률 변수 X 의 값에 따라서 Y 의 확률 분포가 변경된다는 의미에서 **조건부 확률**이라 한다.



선수 지식 - 통계
확률 변수와 확률 분포

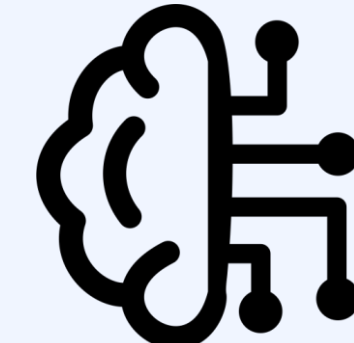
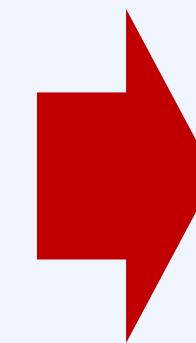
확률질량함수 예시 2)

선수 지식
통계
확률 변수 &
확률 분포

- 한 장의 이미지 x 가 주어졌을 때, 분류 모델의 실행 결과가 다음과 같다고 해보자.
- $P(Y = \text{고양이} | X = x) = 15\%$
- $P(Y = \text{강아지} | X = x) = 55\%$
- $P(Y = \text{다람쥐} | X = x) = 30\%$



이미지 x



딥러닝 모델



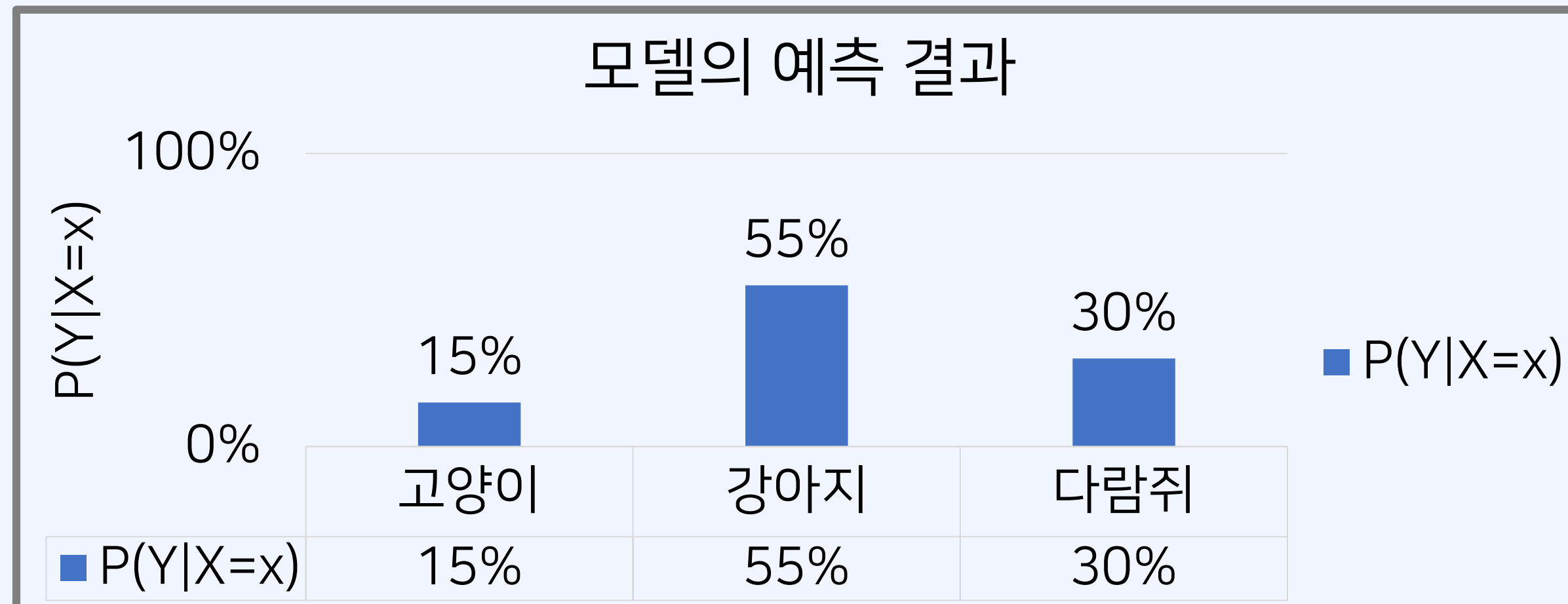
추론 결과

[참고]

- $P(Y|X)$ 는 X 값이 주어졌을 때, 확률 변수 Y 에 대한 **확률 분포**를 의미한다.
- 예시) X 가 이미지, Y 가 클래스(class)라고 하면, 한 장의 이미지가 어떤 동물인지 예측하는 모델의 출력 결과로 이해할 수 있다.

확률질량함수 예시 2)

- 한 장의 이미지 x 가 주어졌을 때, 분류 모델의 실행 결과가 다음과 같다고 해보자.
- $P(Y = \text{고양이} | X = x) = 15\%$
- $P(Y = \text{강아지} | X = x) = 55\%$
- $P(Y = \text{다람쥐} | X = x) = 30\%$
- 즉, 이때 모델을 확률질량함수로 이해할 수 있으며, 확률 분포는 다음과 같이 나타낼 수 있다.



연속확률변수(Continuous Random Variable)

- 확률변수 X 가 취할 수 있는 값이 무한한 경우, 이를 연속확률변수라고 한다.
- 연속적인 값의 예시: 키, 달리기 성적, ...
- 예를 들어 사람의 키는 실수 값을 가지기 때문에, 값의 개수를 셀 수 없다.
- (예시) 키가 정확히 173.32945...일 확률이 얼마일까?
- 이러한 경우 키가 170cm 이상, 175cm 미만일 확률을 구하는 방식을 사용할 수 있다.

확률밀도함수(Probability Density Function, PDF)

- 확률밀도함수(probability density function)란?
- 연속확률변수가 주어진 구간 내에 포함될 확률을 출력하는 함수다.

- 정규 분포(normal distribution) 예시

