확률분포

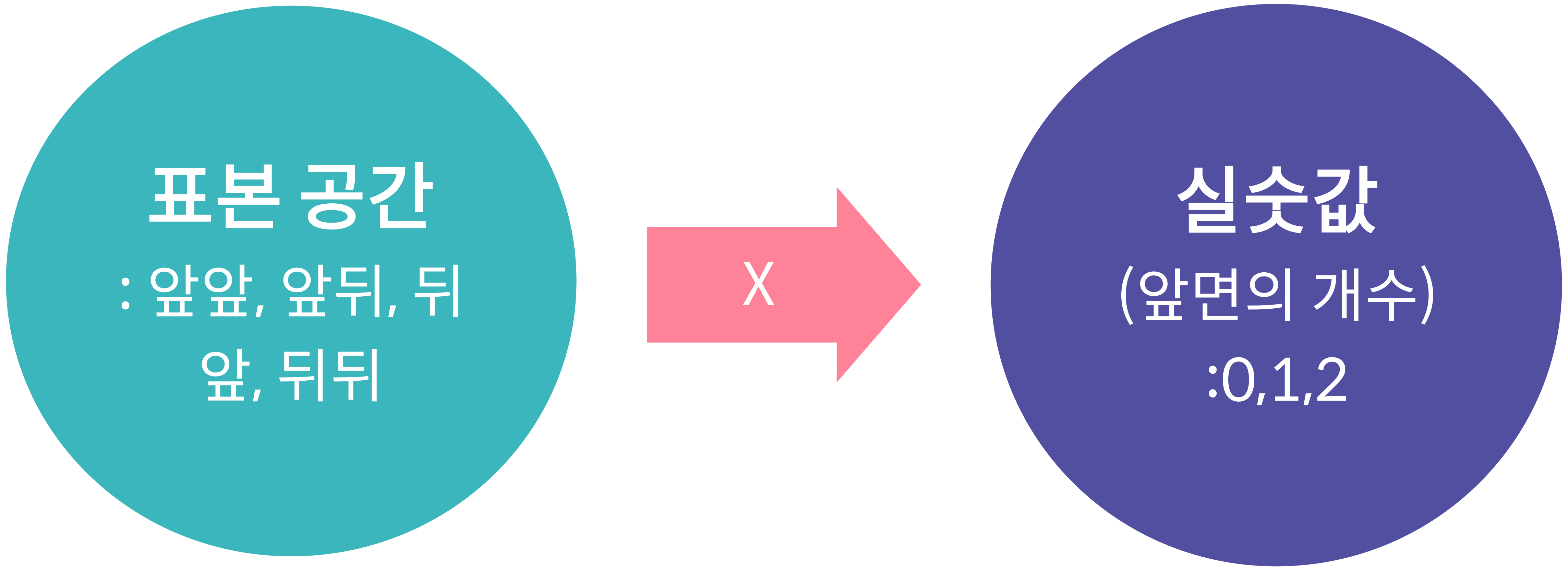
확률 변수

각각의 근원사건에 실수값을 대응시킨 함수

X,Y,⋯처럼 대문자로 표시

시행을 하기 전엔 어떤 값을 갖게 될 지
알 수 없다는 불확실성을 표현

확률 변수



동전을 2번 던졌을 때 앞면의 수를 나타내는 확률변수 X

$$P(X=2) = P(\{앞앞\}) = \frac{1}{4}$$

$$P(X=1) = P(\{앞뒤, 뒤앞\}) = \frac{1}{2}$$

확률 분포

확률 변수가 가질 수 있는 값들이 무엇이며,

그 값을 가질 가능성 또는 확률이

어떻게 분포되어 있는지를

0이상의 실수로 나타낸 것

이산, 연속 확률 변수

이산 확률 변수	연속 확률 변수
확률 변수의 값의 개수를 셀 수 있는 경우	확률 변수의 값이 연속적인 구간에 속하는 경우

이산 확률 분포

변수 x	확률 f(x)
x_1	$f(x_1)$
x_2	$f(x_2)$
$\cdots$	$\cdots$
x_k	$f(x_k)$
합계	1

확률 질량 함수

어떤 확률변수 x가 갖는 확률을 나타내는 함수

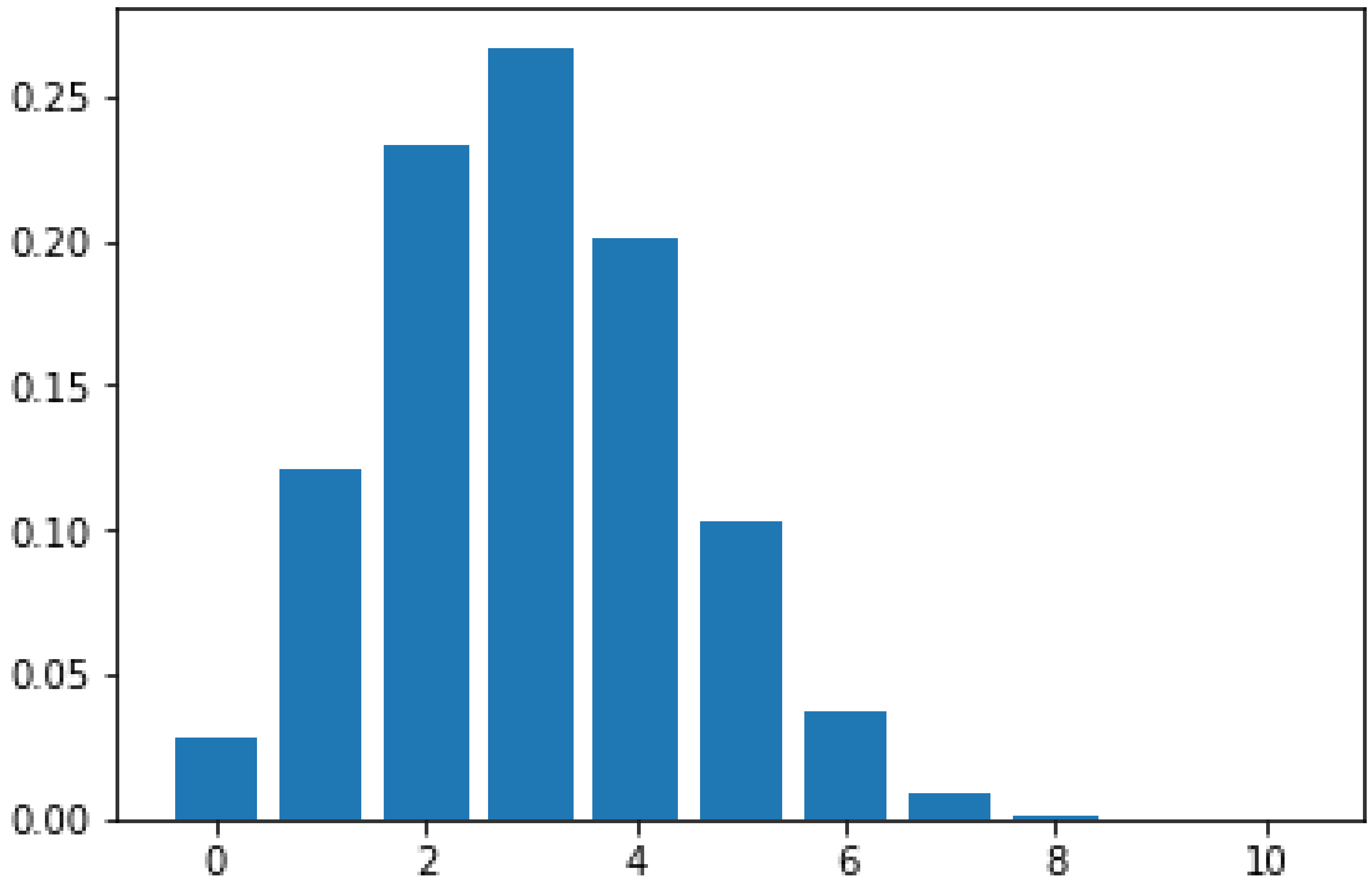
$y = f(x)$: x가 갖는 확률은 y이다

확률 질량 함수의 조건

모든 x_i값에 대해 $0 \leq f(x_i) \leq 1$이고

$\sum_{all\ x_i} f(x_i) = 1$

이산 확률 분포



이산 확률 분포의 종류

1) 베르누이 분포
2) 이항 분포
3) 기하 분포
4) 음이항 분포
5) 초기하 분포
6) 포아송 분포

…

확률 밀도 함수

연속 확률 변수 X가 갖는 확률의 분포를 표현

어느 구간의 확률이 더 크고 작은 지
나타낼 수 있는 함수를 이용

연속 확률 분포

확률 밀도 함수의 조건

1) 모든 x 값에 대해 $f(x) \geq 0$

모든 x값에 대해 확률 밀도 함수값은 0보다 크거나 같다

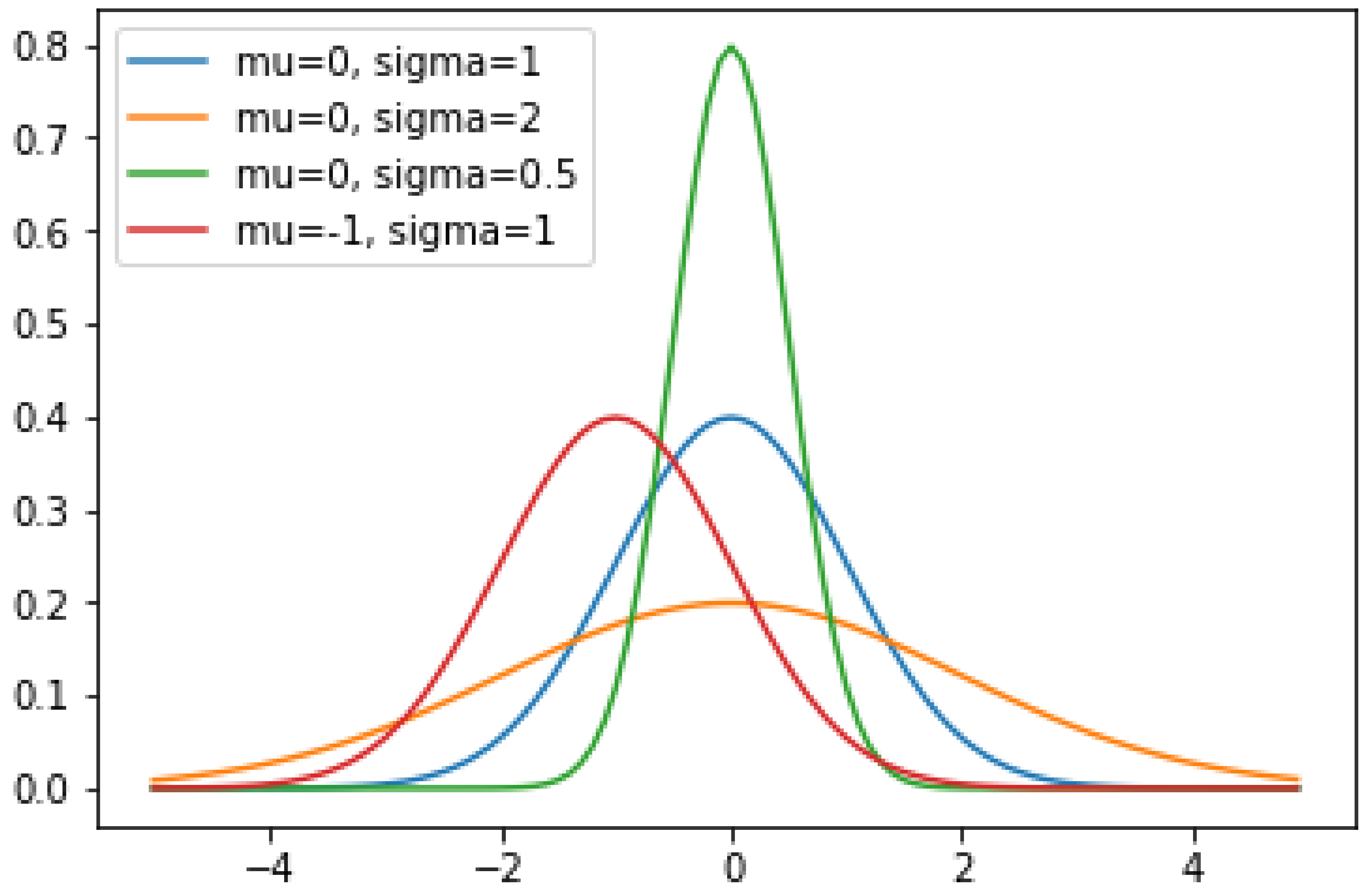
2) $P(a \leq X \leq b) = \int_a^b f(x)dx$

a~b까지 구간의 확률은 그 구간만큼 $f(x)$에서 적분한 값과 같다

3) $P(-\infty < X < \infty) = \int_{-\infty}^{\infty} f(x)dx = 1$

전체 구간을 적분했을 때 확률 밀도 함수값은 1이다

연속 확률 분포


mu=0, sigma=1
mu=0, sigma=2
mu=0, sigma=0.5
mu=-1, sigma=1
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
−4
−2
0
2
4

연속 확률 분포의 종류

1) 균일 분포
2) 지수 분포
3) 감마 분포
4) 정규 분포
5) 베타 분포

…

누적 분포 함수

X가 가질 수 있는 가장 작은 값부터 x까지 해당하는 확률질량함수의 값을 누적해서 더한 것

$F(x) = P(X \leq x)$ 라고 표시

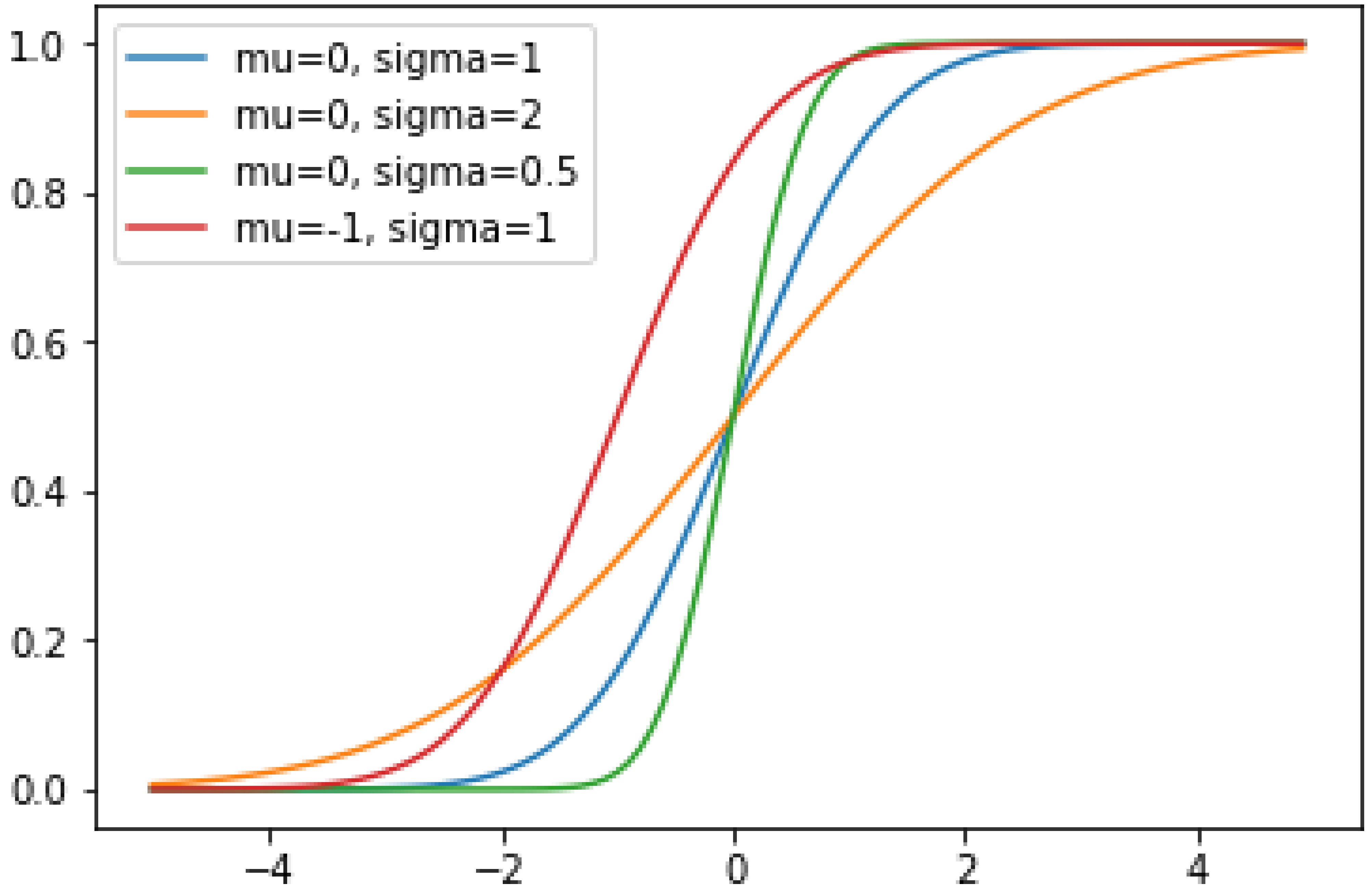
이산확률변수의 누적분포함수

$$F(x) = \sum_{y \leq x, y \epsilon X(s)} p(y)$$

연속확률변수의 누적분포함수

$$F(x) = \int_{-\infty}^{x} f(y)dy$$

연속형 누적 분포 함수



mu=0, sigma=1
mu=0, sigma=2
mu=0, sigma=0.5
mu=-1, sigma=1
1.0
0.8
0.6
0.4
0.2
0.0
−4
−2
0
2
4

이산형 누적 분포 함수