05 강



# 확률분포와 기댓값 1

통계·데이터과학과
이긍희 교수

# 학습목표

1. 이산형 확률변수를 이해할 수 있다.
2. 연속형 확률변수를 이해할 수 있다.
3. 확률질량함수를 이해할 수 있다.
4. 확률밀도함수를 이해할 수 있다.

01

5강 확률분포와 기댓값 1

# 확률변수

# 동전 던지기

- 동전 2개를 던지는 확률실험

$$S = \{(H, H), (H, T)$$
$$(T, H), (T, T)\}$$

# 동전 던지기

◆ 동전 2개 던지기에 따른 상금(X)의 분포

| X | 0 | 100 | 500 | 합 |
|---|---|---|---|---|
| P(X) | 1/4 | 1/2 | 1/4 | 1 |

# 확률변수

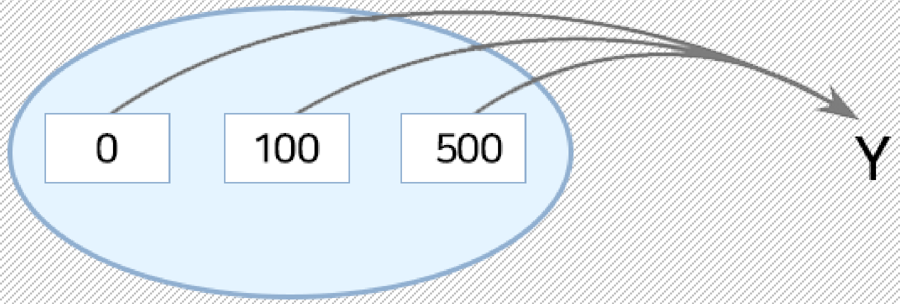
- 확률변수 (random variable) :
  확률적 실험에서 실험결과를 숫자로 표현한 함수

- 이산형 확률변수와 연속형 확률변수

# 이산형 확률변수

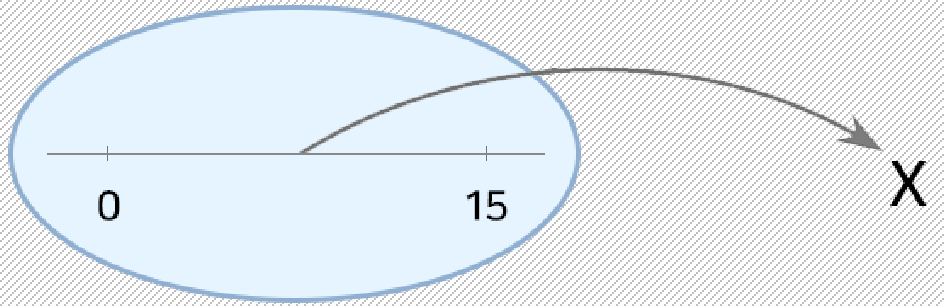
◆ 이산형 확률변수 : 취할 수 있는 값을 셀 수 있을 때

- 확률분포는 점 확률에 의해 결정



# 연속형 확률변수

◆ 연속형 확률변수 : 취할 수 있는 값을 셀 수 없을 때



# 연속형 확률변수

- 연속형 확률변수 : 취할 수 있는 값을 셀 수 없을 때

  - 확률분포는 구간 확률에 의해 결정

# 이산형 확률분포

◆ 이산형 확률분포 : 확률변수가 가질 수 있는 값의 점확률인 확률질량함수에 의하여 확률분포 결정

- 균등분포, 이항분포, 초기하분포, 포아송분포

# 연속형 확률분포

◆ 연속형 확률분포 : 확률변수 각 값의 점확률이 아니라 구간의 확률에 의하여 확률분포가 결정

- 연속형 균등분포, 지수분포, 정규분포

# 누적분포함수

◆ 누적분포함수 F(x) : X가 $(-\infty, x]$에 속할 확률

$$F(x) = P(X \in (-\infty, x]) = P(X \leq x)$$

# 누적분포함수의 성질

- $x \le y \rightarrow F(x) \le F(y)$
- $\lim_{x \to -\infty} F(x) = 0$
- $\lim_{x \to \infty} F(x) = 1$

02

5강 확률분포와 기댓값 1

# 이산형 확률분포

# 이산형 확률변수

## ◆ 이산형 확률변수

취할 수 있는 값이 셀 수 있을 때의 확률변수

(예) 동전의 앞면 수, 불량품 수, 페이지 당 오타 수

# 이산형 확률분포

◆ $S = \{x_1, x_2, \cdots, x_n\}$ , $P(X = x_i) = p_i$

| X | $x_1$ | $x_2$ | $\cdots$ | $x_i$ | $\cdots$ | $x_n$ | 합 |
|---|---|---|---|---|---|---|---|
| P(X) | $p_1$ | $p_2$ | $\cdots$ | $p_i$ | $\cdots$ | $p_n$ | 1 |

# 확률질량함수

◆ **확률질량함수(probability mass function)**

이산형 확률변수의 점확률을 결정지어주는 함수

$$f(x) = P(X = x), \qquad x = x_1, x_2, \cdots, x_n$$

# 확률질량함수의 성질

① $0 \le f(x) \le 1$

② $\sum_{x \in S} f(x) = 1$

③ $P(X \in A) = \sum_{x \in A} f(x)$

# 누적분포함수

- $F(x) = P(X \le x) = \sum_{y \le x} f(y)$
- $f(x) = F(x) - F(x^{-})$

$x_n$

# 이산형 확률분포의 예

**예**

두 개의 주사위를 던져서 나오는 눈금의 합의 확률분포를 구하시오.

# 이산형 확률분포의 예

**예**

두 개의 주사위를 던져서 나오는 눈금의 합의 확률분포를 구하시오.

# 확률질량함수의 예

**예**

두 개의 주사위를 던져서 나오는 눈의 합을 X 라고 할 때 다음을 구하시오.

(1) $P(X = 10)$

| $X$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 합 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $P(X)$ | $\frac{1}{36}$ | $\frac{2}{36}$ | $\frac{3}{36}$ | $\frac{4}{36}$ | $\frac{5}{36}$ | $\frac{6}{36}$ | $\frac{5}{36}$ | $\frac{4}{36}$ | $a$ | $\frac{2}{36}$ | $\frac{1}{36}$ | 1 |

# 확률질량함수의 예(2)

예

(2) $P(X = 2 \text{ 또는 } X = 12)$

| $X$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 합 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $P(X)$ | $\frac{1}{36}$ | $\frac{2}{36}$ | $\frac{3}{36}$ | $\frac{4}{36}$ | $\frac{5}{36}$ | $\frac{6}{36}$ | $\frac{5}{36}$ | $\frac{4}{36}$ | $a$ | $\frac{2}{36}$ | $\frac{1}{36}$ | 1 |

# 확률질량함수의 예

예

(3) $P(X < 7)$

| $X$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 합 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $P(X)$ | $\frac{1}{36}$ | $\frac{2}{36}$ | $\frac{3}{36}$ | $\frac{4}{36}$ | $\frac{5}{36}$ | $\frac{6}{36}$ | $\frac{5}{36}$ | $\frac{4}{36}$ | $a$ | $\frac{2}{36}$ | $\frac{1}{36}$ | 1 |

# 확률질량함수의 예

**예** (3) $P(X \geq 7)$

| $X$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 합 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $P(X)$ | $\frac{1}{36}$ | $\frac{2}{36}$ | $\frac{3}{36}$ | $\frac{4}{36}$ | $\frac{5}{36}$ | $\frac{6}{36}$ | $\frac{5}{36}$ | $\frac{4}{36}$ | $a$ | $\frac{2}{36}$ | $\frac{1}{36}$ | 1 |

# 확률질량함수의 예

예

(4) $P(X < 7$ 또는 $X$는 짝수)

| $X$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 합 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $P(X)$ | $\frac{1}{36}$ | $\frac{2}{36}$ | $\frac{3}{36}$ | $\frac{4}{36}$ | $\frac{5}{36}$ | $\frac{6}{36}$ | $\frac{5}{36}$ | $\frac{4}{36}$ | $a$ | $\frac{2}{36}$ | $\frac{1}{36}$ | 1 |

03

5강 확률분포와 기댓값 1

# 연속형 확률분포

# 연속형 확률변수의 정의

◆ **연속형 확률변수**

어떤 구간에 속하는 연속적인 값을 가지는 확률변수

- (예) 키, 무게, 사용기간, 주가지수

# 누적분포함수

- 확률변수 X가 $(-\infty, x]$에 속할 확률로 정의

$$F(x) = P(X \leq x), -\infty < x < \infty$$

# 연속형 확률분포의 예

**예** 버스가 정류장에 15분 간격으로 도착

(1) 버스를 기다리는 시간이 10분일 확률은?

# 연속형 확률분포의 예

예

버스가 정류장에 15분 간격으로 도착

(2) 버스를 기다리는 시간이 10분보다 짧을 확률은?

# 연속형 확률분포의 예

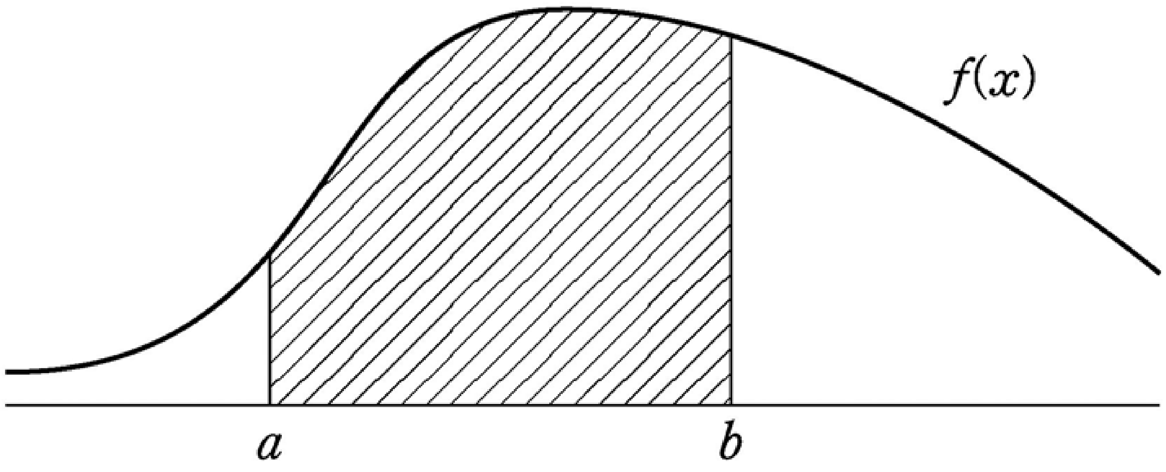
**예**

버스가 정류장에 15분 간격으로 도착

(3) 버스를 기다리는 시간의 누적분포함수는?

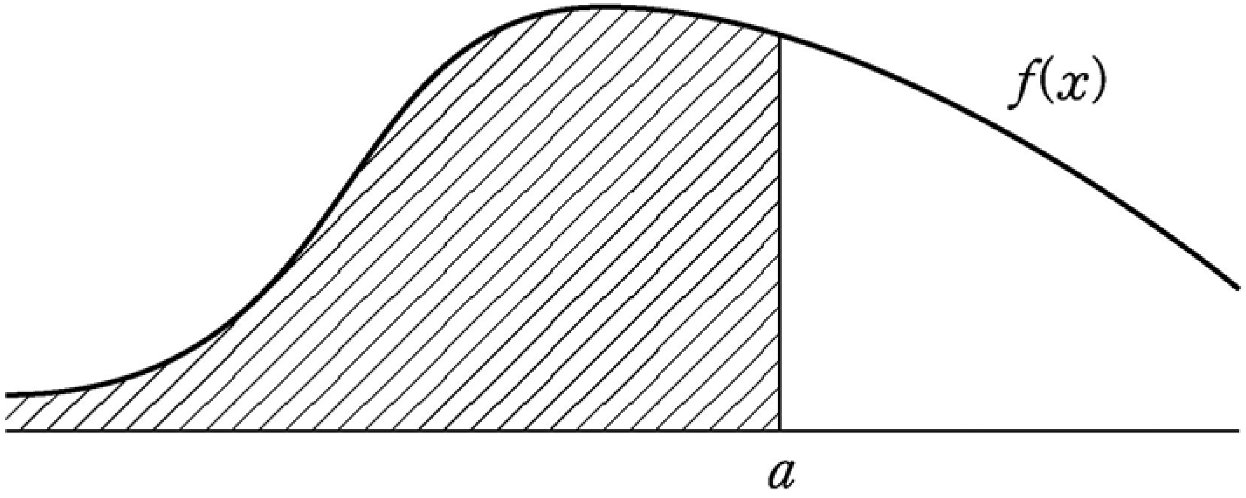
# 확률밀도함수

- 확률밀도함수(probability density function) : 연속형 확률변수가 구간에 속할 확률을 결정지어 주는 함수



# 누적분포함수

- 누적분포함수 F(a): X가 a보다 작을 확률



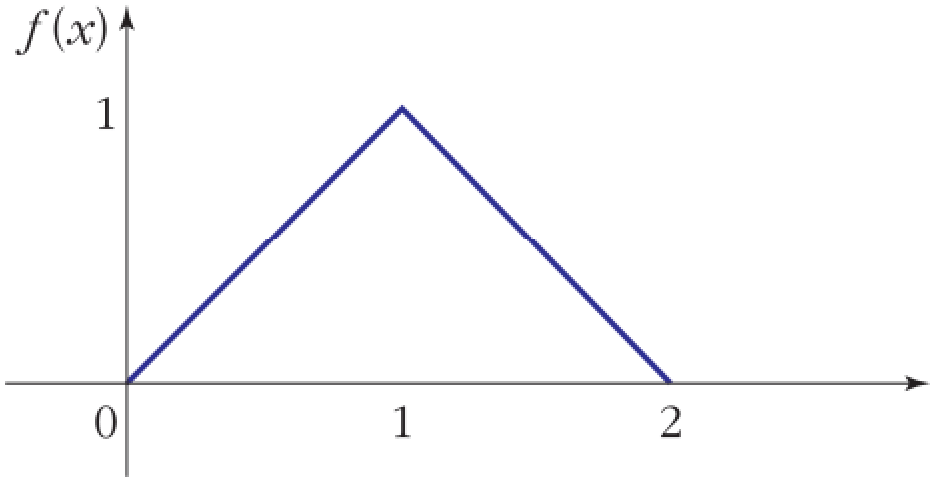
# 누적분포함수

- $F(x) = \int_{-\infty}^{x} f(t)dt$

# 누적분포함수의 예

**예**

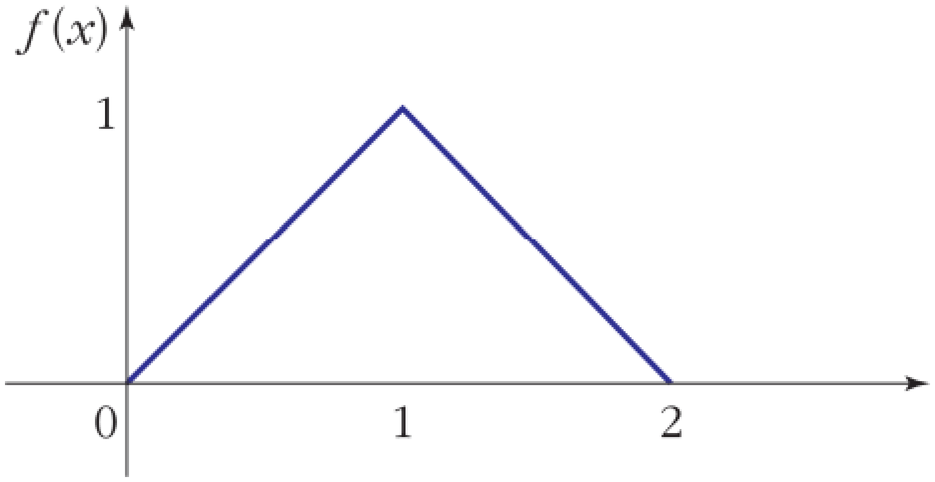
X가 0.5이하일 확률은?



# 누적분포함수의 예

예

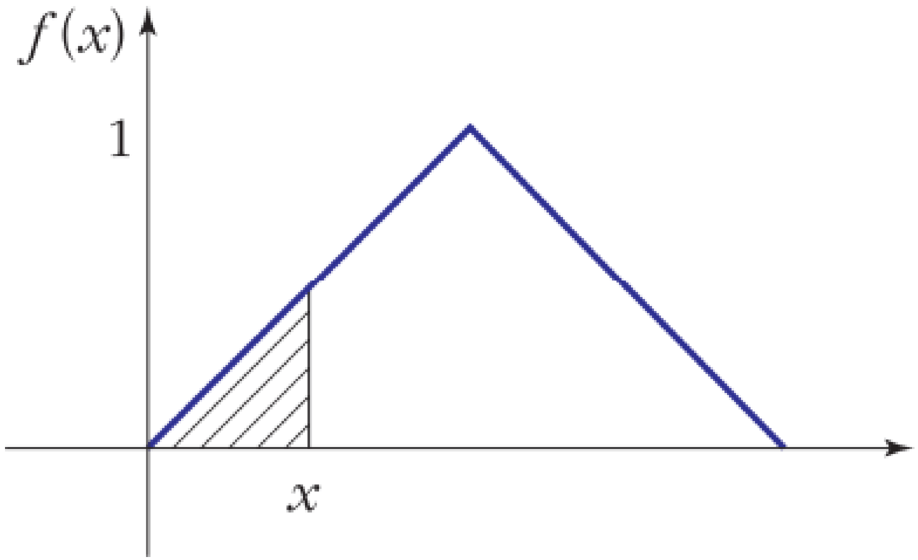
X가 0.5보다 크고 1.5보다 작을 확률은?



# 누적분포함수의 예

**예**

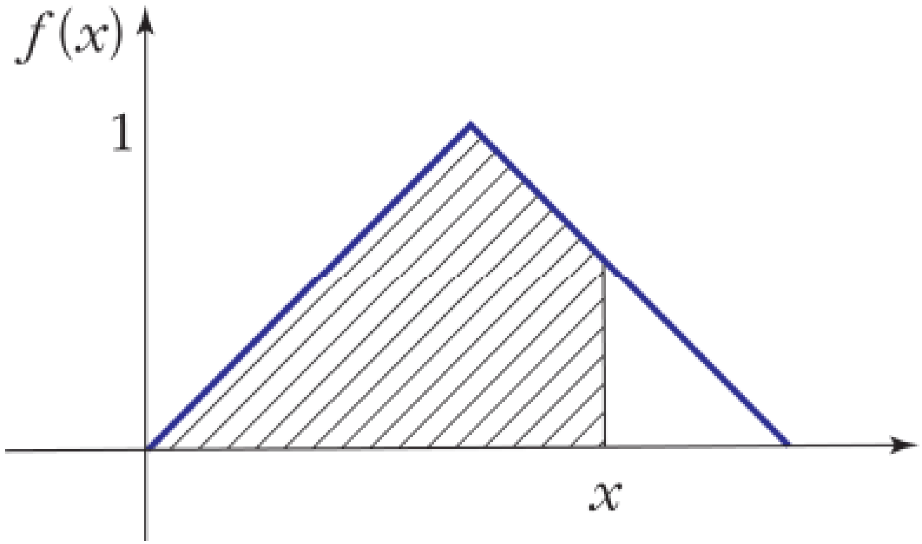
X의 누적분포함수는?



# 누적분포함수의 예

예

X의 누적분포함수는?



# 확률밀도함수의 성질

① $f(x) \geq 0$

② $\int_{-\infty}^{\infty} f(x)\,dx = 1$

③ $P(a \leq X \leq b) = \int_{a}^{b} f(x)\,dx$

# 확률밀도함수의 예

예

$$f(x) = kx^2, \qquad 0 \leq x \leq 2$$

(1) $k$ 값은?

# 확률밀도함수의 예

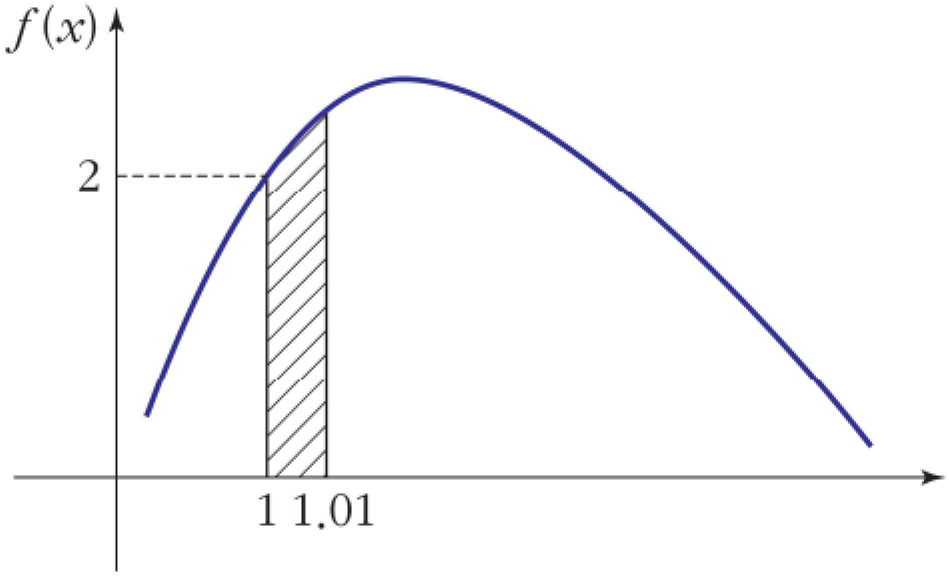
예

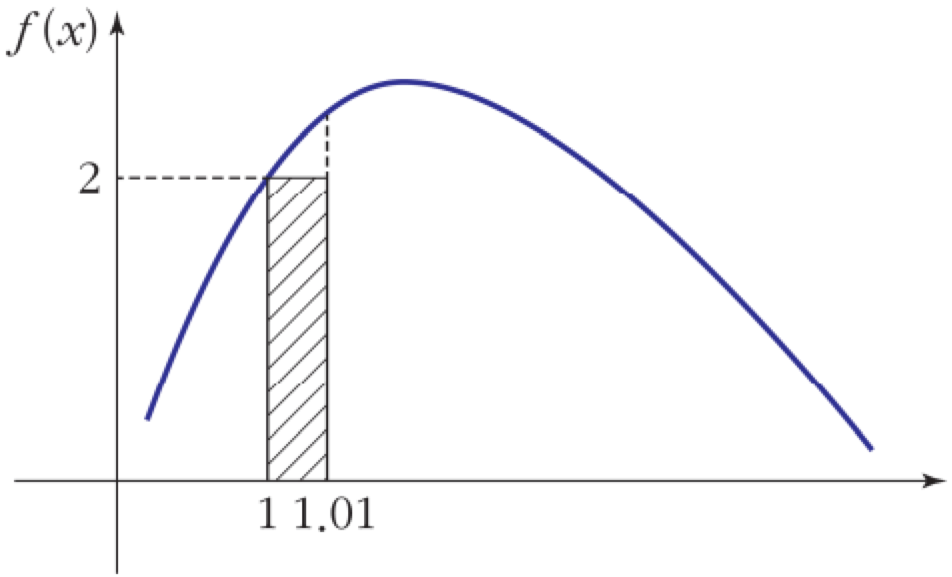
$$f(x) = kx^2, \qquad 0 \le x \le 2$$

(2) X가 1보다 작은 값을 가질 확률은?

# 확률밀도함수

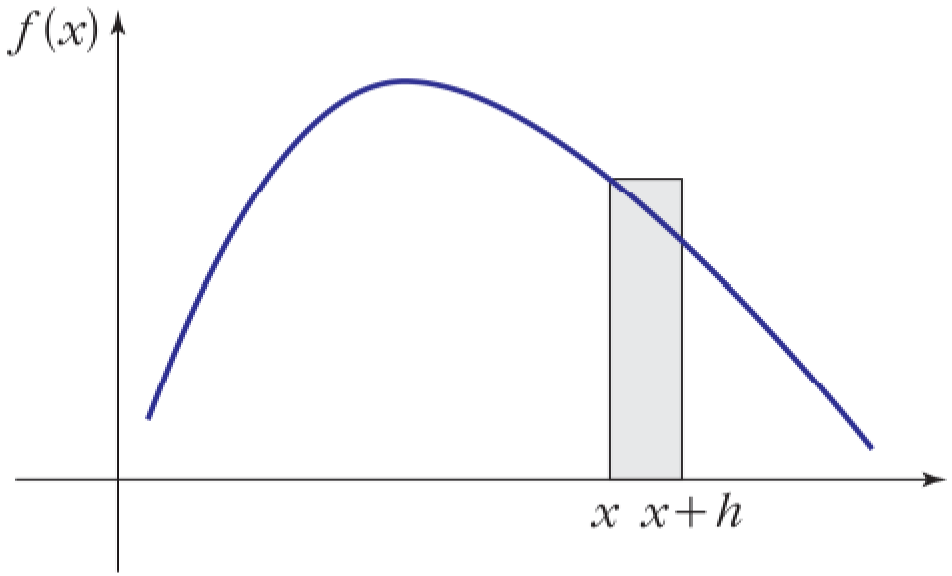
◆ $P(x \le X \le x+h) \approx f(x) \times h$



# 확률밀도함수

- $P(x \le X \le x+h) \approx f(x) \times h$



# 누적분포함수와 확률밀도함수

- $F(x) = \int_{-\infty}^{x} f(t)dt$

- $f(x) = \frac{d}{dx}F(x)$

# 확률밀도함수의 예

**예**

X의 누적분포함수가 다음과 같을 때 X의 확률밀도함수는?

$$F(x) = \begin{cases} 0, & x < 0 \text{일 때} \\ x^2, & 0 \le x \le 1 \text{일 때} \\ 1, & x > 1 \text{일 때} \end{cases}$$

# 학습정리

- 확률적 실험에서 실험결과를 수치로 나타낸 것을 확률변수라고 한다.
- 확률변수가 취할 수 있는 값을 셀 수 있을 때 이를 이산형 확률변수라 하고, 그렇지 않고 연속형일 때 연속형 확률변수라고 한다.

정리하기

# 학습정리

정리하기

- 이산형 확률변수의 확률분포는 확률질량함수에 의하여, 연속형 확률변수의 확률분포는 확률밀도함수에 의하여 결정된다.

- 확률질량함수는 0과 1 사이의 값을 가지며, 모든 경우의 합은 1이 된다. 또한 서로 배반적인 값들의 합집합의 확률은 각각의 확률값의 합으로 표현된다.

# 학습정리

- 확률밀도함수는 0을 포함한 양의 값을 가지며, $x$축과 확률밀도함수로 둘러싸인 부분의 전체 넓이는 1이 된다. 연속형 확률변수가 구간에 속할 확률은 그 구간에서 $x$ 축과 확률밀도함수로 둘러싸인 부분의 넓이이다.

정리하기

다음시간에는

# 수고하셨습니다.

05 강 확률분포와 기댓값 1

06 강 확률분포와 기댓값 2