# 2장. 랜덤변수

2.1 개요
2.2 랜덤변수의 정의
2.3 랜덤변수로 정의되는 사건
2.4 분포함수
2.5 이산랜덤변수
2.6 연속랜덤변수

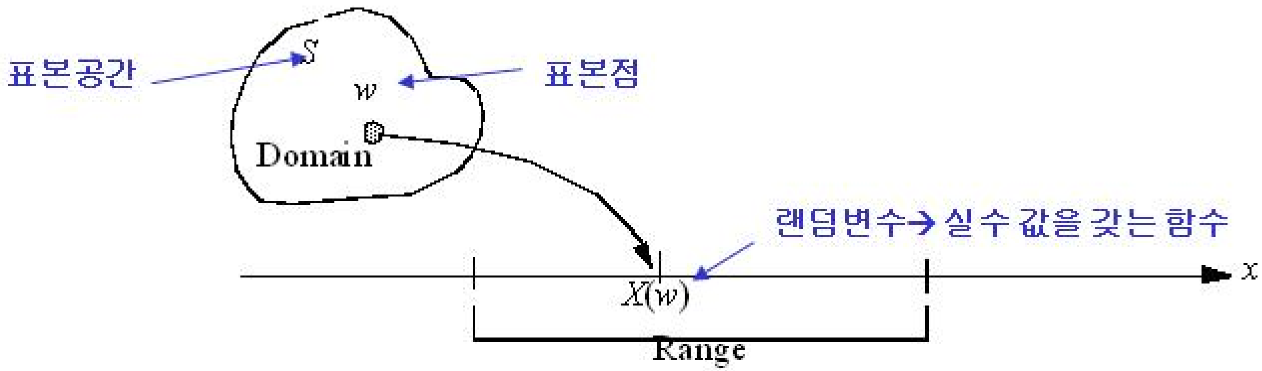
## 2.1 개요

(생략)

## 2.2 랜덤변수의 정의

■ 랜덤변수 (Random Variable)

: 표본공간 $S$의 각 표본점 $w$에 실수 값을 대응시키는 함수



(예) 평범한 동전을 두 개 던지는 실험

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

$X$ = 평범한 동전을 두 개 던지는 실험에서 나온 앞면의 개수

| (사건) | | (랜덤변수) |
|---|---|---|
| $\{TT\}$ | $\Rightarrow$ | $[X=0]$ |
| $\{HT, TH\}$ | $\Rightarrow$ | $[X=1]$ |
| $\{HH\}$ | $\Rightarrow$ | $[X=2]$ |

(참고) $\{TT\}$는 집합으로 표현된 사건이고, $[X=0]$는 랜덤변수로 표현된 사건임

## 2.3 랜덤변수로 정의되는 사건

▶ 랜덤변수로 표현된 사건과 집합으로 표현된 사건

$$[X = x] = \{w \mid X(w) = x\}$$

$$[X \leqq x] = \{w \mid X(w) \leqq x\}$$

$$[X > x] = \{w \mid X(w) > x\}$$

$$[a < X < b] = \{w \mid a < X(w) < b\}$$

(참고) 랜덤변수로 표현된 사건의 확률과 집합으로 표현된 동일한 사건의 확률은 같다.

$$P[X = x] = P\{w \mid X(w) = x\}$$

(예) 평범한 동전을 두 개 던지는 실험

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

$X$ = 평범한 동전을 두 개 던지는 실험에서 나온 앞면의 개수

$$P[X=0] = P\{TT\} = \frac{1}{4}$$

$$P[X=1] = P\{HT, TH\} = \frac{2}{4}$$

$$P[X=2] = P\{HH\} = \frac{1}{4}$$

$$\sum_{n=0}^{2} P[X=n] = 1$$

$$P[X \leqq 1] = P\{HT, TH, TT\} = \frac{3}{4}$$

(예) 앞면이 나올 확률이 $p$ 인 동전을 앞면이 나올 때까지 던지는 실험을 한다.

$$S = \{H, TH, TTH, TTTH, \cdots\}$$

$N =$ 앞면이 나올 때까지 동전을 던진 횟수

$$P[N=1] = P\{H\} = p$$
$$P[N=2] = P\{TH\} = (1-p)p$$
$$P[N=3] = P\{TTH\} = (1-p)^2 p$$
$$P[N=4] = P\{TTTH\} = (1-p)^3 p$$
$$\vdots$$
$$P[N=n] = P\{TT\cdots TH\} = (1-p)^{n-1} p$$

$$\sum_{n=1}^{\infty} P[X=n] = p\sum_{n=1}^{\infty} (1-p)^{n-1}$$
$$= p\frac{1}{1-(1-p)} = 1$$

## ■ 랜덤변수

(1) 이산랜덤변수 (Discrete Random Variable)

: 랜덤변수가 유한하거나, 또는 셀 수 있는 값들을 갖는다.

(2) 연속랜덤변수 (Continuous Random Variable)

: 랜덤변수가 어떤 구간에 속한 연속적인 실수 값들을 갖는다.

## 2.4 분포함수

■ 누적분포함수 (cdf: cumulative distribution function)

$$F_X(x) = P[X \leqq x]$$

▶ 누적분포함수의 특성

(1) $F_X(x)$는 감소하지 않는 함수 ($x_1 < x_2$이면, $F_X(x_1) \leqq F_X(x_2)$)

(2) $0 \leqq F_X(x) \leqq 1$

(3) $F_X(\infty) = 1$

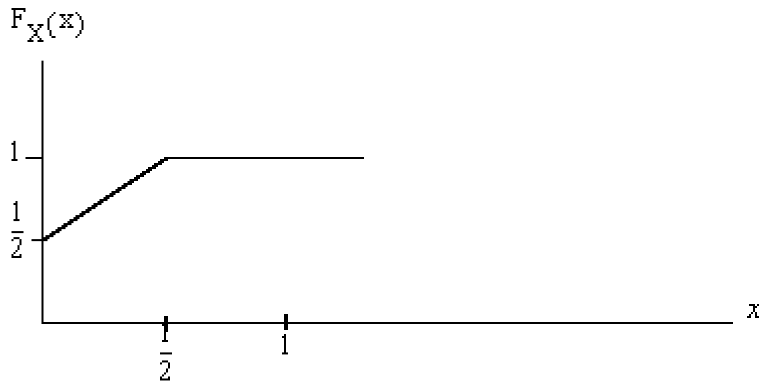
(4) $F_X(-\infty) = 0$

(5) $P[a < X \leqq b] = F_X(b) - F_X(a)$

(6) $P[X > a] = 1 - P[X \leqq a] = 1 - F_X(a)$

(예) 랜덤변수 $X$에 대한 누적분포함수가 다음과 같다.

$$F_X(x) = \begin{cases} 0 & x < 0 \\ x + \dfrac{1}{2} & 0 \leq x \leq \dfrac{1}{2} \\ 1 & x > \dfrac{1}{2} \end{cases}$$




(a) 누적분포함수의 그래프를 그려라

(b) $P[X > \frac{1}{4}]$을 구하라.

$$P[X > \frac{1}{4}] = 1 - P[X \leq \frac{1}{4}] = 1 - F_X(\frac{1}{4}) = 1 - (\frac{1}{4} + \frac{1}{2}) = \frac{1}{4}$$

## 2.5 이산랜덤변수

### ■ 확률질량함수 (pmf: probability mass function)

$$p_X(x) = P[X = x]$$

▶ 확률질량함수의 특성

(1) 만약 $X$가 $x_1, x_2, \cdots, x_n$ 값들만을 갖는 이산랜덤변수이면,

$$p_X(x_i) > 0, \quad i = 1, 2, \cdots, n$$

$$p_X(x_i) = 0, \quad \text{그 외 } i$$

(2) $\displaystyle\sum_x p_X(x) = 1$

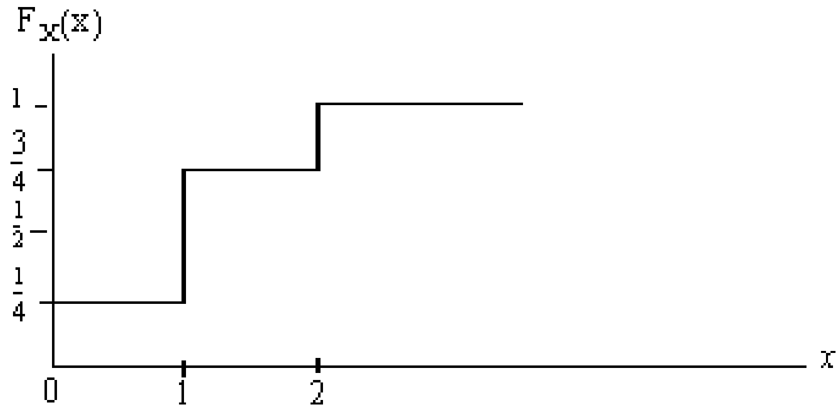
(3) $\displaystyle F_X(x) = \sum_{k \leqq x} p_X(k)$

(예) $X$의 확률질량함수(pmf)가 다음과 같다.

$$p_X(0) = \frac{1}{4},\ p_X(1) = \frac{1}{2},\ p_X(2) = \frac{1}{4}$$

이 때 $X$의 누적분포함수는 다음과 같다.

$$F_X(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{4} & 0 \leqq x < 1 \\ \frac{3}{4} & 1 \leqq x < 2 \\ 1 & x \geqq 2 \end{cases}$$



(예) $X$ = 평범한 동전을 3회 던졌을 때 나오는 앞면의 개수

$$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

$X$의 확률질량함수

$$p_X(0) = P[X = 0] = P\{TTT\} = \frac{1}{8}$$

$$p_X(1) = P[X = 1] = P\{HTT, THT, TTH\} = \frac{3}{8}$$

$$p_X(2) = P[X = 2] = P\{HHT, HTH, THH\} = \frac{3}{8}$$

$$p_X(3) = P[X = 3] = P\{HHH\} = \frac{1}{8}$$

(예) 이산랜덤변수 $X$의 누적분포함수가 다음과 같을 때 확률질량함수를 구하라.

$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1/6 & 0 \leqq x < 2 \\ 1/2 & 2 \leqq x < 4 \\ 5/8 & 4 \leqq x < 6 \\ 1 & x \geqq 6 \end{cases}$$

$$p_X(0) = P[X = 0] = \frac{1}{6} - 0 = \frac{1}{6}$$

$$p_X(2) = P[X = 2] = \frac{1}{2} - \frac{1}{6} = \frac{1}{3}$$

$$p_X(4) = P[X = 4] = \frac{5}{8} - \frac{1}{2} = \frac{1}{8}$$

$$p_X(6) = P[X = 6] = 1 - \frac{5}{8} = \frac{3}{8}$$

(참고) $p_X(0) + p_X(2) + p_X(4) + p_X(6) = 1$

## 2.6 연속랜덤변수

■ 확률밀도함수 (pdf: probability density function)

: 연속랜덤변수 $X$는 모든 실수 $x \in (-\infty, \infty)$에 대하여 정의되고, 임의의 실수 집합 $A$에 대하여 다음과 같은 성질을 갖는 음이 아닌 함수 (확률밀도함수) $f_X(x)$가 존재한다.

$$P[X \in A] = \int_A f_X(x)\,dx$$

▶ 확률밀도함수의 특성

(1) $f_X(x) > 0$

(2) $\int_{-\infty}^{\infty} f_X(x)\,dx = P[-\infty < X < \infty] = 1$

(3) $P[a \leqq X \leqq b] = \int_{a}^{b} f_X(x)\,dx$

(4) $P[X = a] = \int_{a}^{a} f_X(x)\,dx = 0$

(연속랜덤변수는 특정 실수 값을 가질 확률이 0이다.)

(5) $F_X(a) = \int_{-\infty}^{a} f_X(x)\,dx$

($F_X(a) = P[X \leqq a] = P[X < a]$)

(6) $f_X(x) = \frac{d}{dx} F_X(x)$

(예) 연속랜덤변수 $X$의 확률밀도함수가 다음과 같다.

$$f_X(x) = \begin{cases} A(2x - x^2) & 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

(a) $A$값을 구하라.

$$\int_{-\infty}^{\infty} f_X(x)\,dx = 1\text{에서}$$

$$\int_0^2 f_X(x)\,dx = \int_0^2 A(2x - x^2)\,dx = A\left[x^2 - \frac{1}{3}x^3\right]_0^2 = \frac{4}{3}A = 1$$

$$\therefore A = \frac{3}{4}$$

(b) $P[X > 1]$을 구하라.

$$P[X > 1] = \int_0^1 f_X(x)\,dx = \int_0^1 \frac{3}{4}(2x - x^2)\,dx$$

$$= \frac{3}{4}\left[x^2 - \frac{1}{3}x^3\right]_0^1 = \frac{3}{4} \cdot \frac{2}{3} = \frac{1}{2}$$

(예) 연속랜덤변수 $X$의 확률밀도함수가 다음과 같다.

$$g_X(x) = \begin{cases} c & a \leqq x \leqq b \\ 0 & \text{otherwise} \end{cases}$$

(a) $g_X(x)$가 타당한 확률밀도함수가 되기 위한 c 값은?

$$\int_{-\infty}^{\infty} g_X(x)\,dx = 1 \text{에서}$$

$$\int_a^b g_X(x)\,dx = \int_a^b c\,dx = [cx]_a^b = c(b-a) = 1 \qquad \therefore c = \frac{1}{b-a}$$

(b) 랜덤변수 $X$의 누적분포함수를 구하라.

$$G_X(x) = \int_{-\infty}^{x} g_X(u)\,du = \int_a^x \frac{1}{b-a}\,du$$

$$= \begin{cases} 0 & x < a \\ \dfrac{x-a}{b-a} & a \leqq x < b \\ 1 & x \geqq b \end{cases} \qquad \left(\int_a^x \frac{1}{b-a}\,du = \frac{x-a}{b-a}\right)$$

(예) 연속랜덤변수 $X$의 누적분포함수가 다음과 같다.

$$F_X(x) = \begin{cases} 0 & x < 2 \\ A(x-2) & 2 \leqq x < 6 \\ 1 & x \geqq 6 \end{cases}$$

(a) $X$의 확률밀도함수를 구하라.

$$f_X(x) = \frac{d}{dx} F_X(x) = \begin{cases} A & 2 \leqq x \leqq 6 \\ 0 & \text{otherwise} \end{cases}$$

(b) $A$ 값을 구하라.

$\int_{-\infty}^{\infty} f_X(x)\,dx = 1$에서,

$$\int_2^6 f_X(x)\,dx = \int_2^6 A\,dx = [Ax]_2^6 = 4A = 1 \qquad \therefore A = \frac{1}{4}$$

(c) $P[X > 4]$을 구하라.

$$P[X > 4] = \int_4^6 f_X(x)\,dx = \int_4^6 \frac{1}{4}\,dx = \frac{1}{2}$$

(연습문제 2장) 1, 2, 7, 16, 22, 23, 25, 31