4주차: 확률변수와 분포(이산확률변수와 연속확률변수 정의, 확률질량함수와 확률밀도함수 정의)

4.1 확률변수와 누적분포함수

정의

확률변수 = 정의역이 표본공간 S 이고 공역이 실수 전체 집합 R인 함수

기호로 영어 대문자($X, Y, Z, \cdots$) 로 나타냄.

이산 확률변수 = 확률변수의 치역의 원소의 개수를 셀 수 있는 확률변수

연속 확률변수 = 확률변수의 치역의 원소의 개수를 셀 수 없는 확률변수

[예]

(1)50명의 학생으로 구성된 한 학급의 수학성적을 X라고 했을 때

X의 치역은 많아야 50개의 성적이므로

X의 치역의 원소의 개수를 셀 수 있으므로 X는 이산확률변수.

(2) 구간 $(0,1)$의 모든 실수의 값을 가지는 확률변수 X는 연속 확률변수이다.

X의 치역의 원소의 개수를 셀 수 없기 때문이다.

정의

누적분포함수 = 확률변수에 대하여 누적된 확률

기호로 $F(x) = P(X \le x)$

[연습문제] 다음 물음에 답하여라.

어떤 시행의 확률변수 X 에 대하여 실험이 성공하면 $X=1$ 이고 실험이 실패하면 $X=0$ 으로 주어졌다고 하자. X는 이산확률변수인가?

4.2 이산확률변수와 확률질량함수

정의

이산확률변수 X의 확률질량함수(probabilty mass function ; pmf) = $P(X=x_i)$ 또는 $f(x_i)$

기호: $p_i = P(X=x_i)$

이산확률분포표

X	x_1	x_2	$\cdots$	x_n	합
P	p_1	p_2	$\cdots$	p_n	1

[예]두 개의 동전을 던지는 시행에서 앞면이 나오는 동전의 개수를 X 라고 할 때, X 의 확률분포표를 구하여라.

풀이:

X	0	1	2	합
P	1/4	1/2	1/4	1

[예] 어떤 시행의 확률변수 X 에 대하여 실험이 성공하면 $X=1$ 이고 실험이 실패하면 $X=0$ 으로 주어졌다고 하자. 실험이 성공할 확률을 1/3라 할 때, X 의 확률분포표를 구하여라.

풀이:

X	0	1	합
P	2/3	1/3	1

정리

X 가 취하는 값은 $x_i(1 \le i \le n)$이고

X 의 확률질량함수 $f(x_i)$ 일 때, 다음 성립

(1) $0 \le f(x_i) \le 1$

(2) $\sum_{i=1}^{n} f(x_i) = 1$

(3) $P(X \le x_k) = \sum_{i=1}^{k} P(X = x_i) = \sum_{i=1}^{k} f(x_i)$

증명:

확률질량함수와 확률의 정의에 의하여 성립

정리

이산 확률변수 X 의 확률질량함수 $f(x)$ 에 대하여 다음 성립

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

(2) $P(a < X \le b) = F(b) - F(a)$

(3) $P(a \le X \le b) = F(b) - F(a) + f(a)$

증명:

(1)확률질량함수와 확률의 정의에 의하여 성립.

(2) $P(a < X \le b) = P(X \le b) - P(X \le a) = F(b) - F(a)$

(3) $P(a \le X \le b) = P(a < X \le b) + P(X = a)$
$= P(X \le b) - P(X \le a) + P(X = a) = F(b) - F(a) + f(a)$

[연습문제] 다음 물음에 답하여라.

(1) $P(a < X < b) = F(b) - F(a) - f(b)$

(2) $a < b \Rightarrow F(a) \le F(b)$

(3)두 개의 동전을 던지는 시행에서 앞면이 나오는 동전의 개수를 X 라고 할 때,

$F(0), F(1), F(2), F(3)$의 값을 각각 구하여라.

[연습문제] 하나의 동전을 앞면이 나올 때까지 던지는 시행에서 던진 횟수를 X라 할 때, 다음 물음에 답하여라.

(1) 표본공간을 구하여라.

(2) 확률변수 X의 치역을 구하여라.

(3) 앞면이 나올 확률을 p라 할 때, $\mathrm{P}(X=2)$를 구하여라.

(4) $\mathrm{P}(X=n)$을 구하여라.

(5) $\mathrm{P}(\bigcup_{n=1}^{\infty}\{X=n\})$을 구하여라.

(6) X의 누적분포함수를 구하여라.

4.3 연속확률변수와 확률밀도함수

정의

연속확률변수 X의 확률밀도함수 $f(x)$ if

$$P(X \le x) = \int_{-\infty}^{x} f(t)\, dt$$

확률밀도함수$f(x)$는 누적분포함수 $F(x)$를 이용하여 정의되었고

누적분포함수를 미분하면 확률밀도함수를 구할 수 있다.

즉, $f(x) = F'(x)$

확률밀도함수(probability density function; pdf)는 다음 성질을 만족

정리

X 의 pdf $f(x)$ 에 대하여, 다음 성립

(1) 실수 x에 대하여 $0 \le f(x)$

(2) $\int_{-\infty}^{\infty} f(t)\, dt = 1$

(3)실수 a에 대하여 $P(X=a)=0$

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

증명:

(1) $F(x) = \int_{-\infty}^{x} f(t)\, dt$ 를 미분하면 $F'(x) = f(x)$ 이고

F는 증가함수이므로 도함수 $F'(x) \ge 0$ 이다.

따라서 $f(x) = F'(x) \ge 0.$

(2) $\int_{-\infty}^{\infty} f(t)\, dt = \lim_{x \to \infty} P(X \le x) = P(X < \infty) = 1$

(3) $P(X=a) = P(a \le X \le a) = \int_{a}^{a} f(t)\, dt = 0$

(4) $P(a \le X \le b) = P(X \le b) - P(X \le a) + P(X=a) = P(X \le b) - P(X \le a)$

$$= \int_{-\infty}^{b} f(t)\, dt - \int_{-\infty}^{a} f(t)\, dt = \int_{-\infty}^{b} f(t)\, dt - \int_{a}^{-\infty} f(t)\, dt$$

$$= \int_{a}^{b} f(t)\, dt$$

[예] X 의 확률밀도함수가

$$f(x)=\begin{cases} k & (0 \leqq x \leqq 2) \\ 0 & (x<0, \ x>2) \end{cases}$$

일 때, 다음을 구하여라.

(1)상수 k

(2) $P(0 \leq X \leq 1)$

(3) $P(X \leq 2)$

풀이:

(1) $\int_{-\infty}^{\infty} f(t)\,dt = 1$ 이므로 $1 = \int_{-\infty}^{\infty} f(t)\,dt = \int_{0}^{2} k\,dt$ 이다. 따라서 $k = \frac{1}{2}$

(2) $P(0 \leq X \leq 1) = \int_{0}^{1} f(t)\,dt = \int_{0}^{1} \frac{1}{2}\,dt = \frac{1}{2}$

(3) $P(X \leq 2)$에서 사건 $X \leq 2$는 전사건이므로 $P(X \leq 2) = 1$

[연습문제]

임의의 실수 $\lambda > 0$에 대하여 확률밀도함수

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & (0 \leqq x < \infty) \\ 0 & (x < 0) \end{cases}$$

를 가지는 연속확률변수 X에 대하여 다음을 구하여라.

(1) $P(0 \leq X \leq 1)$

(2) $P(X \leq 2)$

(3)누적분포함수를 구하여라.