Ⅲ_1. 이산확률변수의 확률분포

[12확통03-01] 확률변수와 확률분포의 뜻을 안다.
[12확통03-02] 이산확률변수의 기댓값(평균)과 표준편차를 구할 수 있다.
[12확통03-03] 이항분포의 뜻을 알고, 평균과 표준편차를 구할 수 있다.

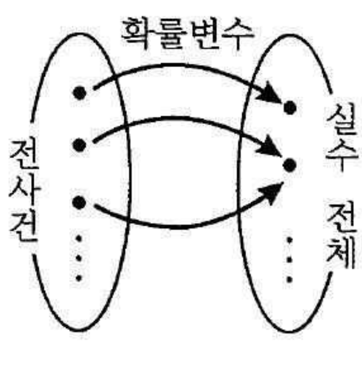
1 확률변수 ❶

(1) 확률변수

어떤 시행에서 표본공간의 각 원소에 하나의 실수의 값을 대응시키는 함수를 '확률변수'라고 한다. 확률변수 X가 어떤 값 x를 가질 확률을 기호로 $\mathrm{P}(X=x)$와 같이 나타낸다.

예 세 개의 동전을 동시에 던지는 시행에서 앞면이 나오는 동전의 개수를 확률변수 X라 하면 X가 갖는 값은 0, 1, 2, 3이다.

☑ 확률변수는 표본공간을 정의역으로 하고 실수 전체의 집합을 공역으로 하는 함수이지만 변수의 역할도 하기 때문에 확률변수라고 한다.

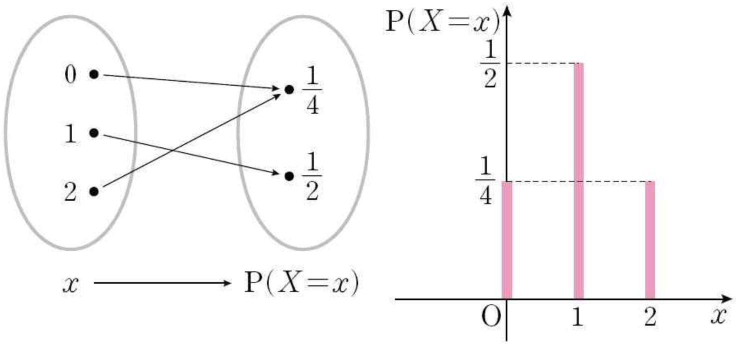


1 확률변수 ❷

(2) 이산확률변수

확률변수 X가 갖는 값이 유한개이거나 무한히 많더라도 자연수와 같이 셀 수 있을 때, 그 확률변수 X를 '이산확률변수'라고 한다.

예 두 개의 주사위를 동시에 던지는 시행에서 나온 두 눈의 수의 차를 X라 할 때, X가 갖는 값은 0, 1, 2, 3, 4, 5로 유한개이므로 X는 이산확률변수이다.

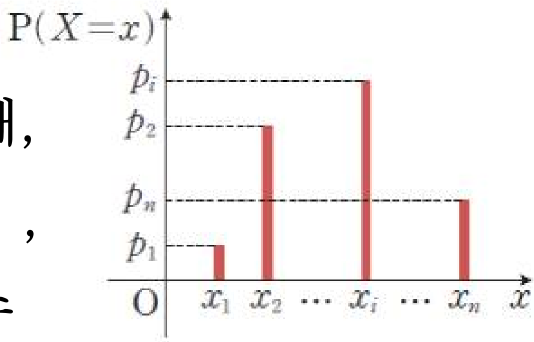


2 이산확률변수의 확률분포 ❶

(1) 이산확률변수의 확률분포

이산확률변수 X가 갖는 값이 x_1, x_2, x_3, $\cdots$, x_n이고 X가 이 값을 가질 확률이 각각 p_1, p_2, p_3, $\cdots$, p_n일 때, x_1, x_2, x_3, $\cdots$, x_n과 p_1, p_2, p_3, $\cdots$, p_n 사이의 대응 관계를 이산확률변수 X의 '확률분포'라고 한다.

X	x_1	x_2	$\cdots$	x_i	$\cdots$	x_n	합계
$\mathrm{P}(X=x_i)$	p_1	p_2	$\cdots$	p_i	$\cdots$	p_n	1



이때 이산확률변수 X의 확률분포는 오른쪽과 같이 표 또는 그래프로 나타낼 수 있다.

2 이산확률변수의 확률분포 ❷

(2) 확률질량함수

이산확률변수 X가 갖는 값 x_i $(i=1,\ 2,\ 3,\ \cdots,\ n)$과 X가 이 값을 가질 확률 p_i $(i=1,\ 2,\ 3,\ \cdots,\ n)$ 사이의 대응 관계를 나타내는 함수

$$\mathrm{P}(X=x_i)=p_i\ \ (i=1,\ 2,\ 3,\ \cdots,\ n)$$

을 이산확률변수 X의 '확률질량함수'라고 한다.

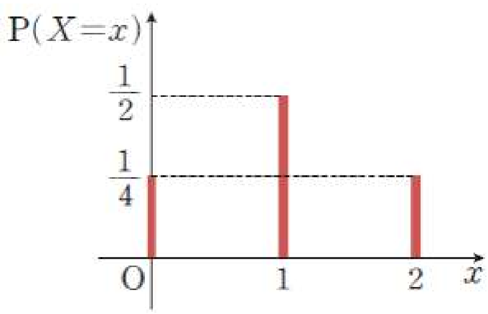
2 이산확률변수의 확률분포 ❸

예 한 개의 주사위를 두 번 던지는 시행에서 4의 약수의 눈이 나온 횟수를 확률변수 X라 하면 X가 갖는 값은 0, 1, 2이므로 X는 이산확률변수이고, X의 확률질량함수는

$$\mathrm{P}(X=x)={}_2\mathrm{C}_x\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{2-x}=\frac{{}_2\mathrm{C}_x}{4}\ \ (x=0,\ 1,\ 2)$$

이다. 이때 확률변수 X의 확률분포를 표와 그래프로 나타내면 오른쪽과 같다.

X	0	1	2	합계
$\mathrm{P}(X=x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	1



$\boxed{3}$ 확률질량함수의 성질 ❶

이산확률변수 X의 확률질량함수가 $\mathrm{P}\left(X=x_i\right)=p_i$ ($i=1$, 2, 3, $\cdots$, n)일 때, 확률의 기본 성질에 의하여 다음이 성립한다.

(1) $0 \le p_i \le 1$

(2) $p_1+p_2+p_3+\cdots+p_n=\sum_{i=1}^{n} p_i=1$

(3) $\mathrm{P}\left(x_i \le X \le x_j\right)=p_i+p_{i+1}+p_{i+2}+\cdots+p_j=\sum_{k=i}^{j} p_k$

(단, $j=1$, 2, $\cdots$, n이고 $i \le j$)

$\boxed{3}$ 확률질량함수의 성질 ❷

예 흰 공 2개와 검은 공 3개가 들어 있는 주머니에서 임의로 3개의 공을 동시에 꺼내는 시행에서 꺼낸 공 중 흰 공의 개수를 확률변수 X라 하면 X가 갖는 값이 0, 1, 2이고 X의 확률질량함수는

$$\mathrm{P}(X=x)=\frac{{}_2\mathrm{C}_x \times {}_3\mathrm{C}_{3-x}}{{}_5\mathrm{C}_3} \quad (x=0,\ 1,\ 2)$$ 이므로

$$\mathrm{P}(X=0)=\frac{1}{10},\ \mathrm{P}(X=1)=\frac{3}{5},\ \mathrm{P}(X=2)=\frac{3}{10}$$

$\therefore\ 0 \le \mathrm{P}(X=x) \le 1$

$\therefore\ \mathrm{P}(X=0)+\mathrm{P}(X=1)+\mathrm{P}(X=2)=1$

이므로 확률질량함수가 위의 성질 (1), (2)를 만족시킨다.

✡ 이산확률변수의 여러 가지 확률분포와 확률질량함수 ❶

(1) 이항분포(Binomial distribution)

$$\mathrm{P}(X=x)={}_{n}\mathrm{C}_{x}\,p^{x}q^{n-x}\quad(x=0,\ 1,\ 2,\ 3,\ \cdots,\ n)$$

(2) 기하분포(Geometric distribution)

$$\mathrm{P}(X=x)=pq^{x-1}$$

$$(x=0,\ 1,\ 2,\ 3,\ \cdots \text{ 이고 } 0<p<1,\ q=1-p)$$

(3) 초기하분포(Hypergeometric distribution)

$$\mathrm{P}(X=x)=\frac{{}_{M}\mathrm{C}_{x}\times{}_{N-M}\mathrm{C}_{x}}{{}_{N}\mathrm{C}_{x}}$$

$$(x=0,\ 1,\ 2,\ 3,\ \cdots,\ n)$$

✡ 이산확률변수의 여러 가지 확률분포와 확률질량함수 ❷

(4) 포아송분포(Poisson distribution)

$$\mathrm{P}(X=x)=\frac{e^{-\mu}\mu^{x}}{x!}\quad(x=0,\ 1,\ 2,\ 3,\ \cdots)$$

(5) 이산균등분포(Discontinuous distribution)

$$\mathrm{P}(X=x)=\frac{1}{n}\quad(x=0,\ 1,\ 2,\ 3,\ \cdots,\ n)$$

(6) 음의 이항분포(Negative Binomial distribution)

$$\mathrm{P}(X=x)={}_{x-1}\mathrm{C}_{k-1}\,p^{k}q^{x-k}$$

$$(x=k,\ k+1,\ k+2,\ \cdots)$$

4 이산확률변수 X의 기댓값(평균) ❶

이산확률변수 X의 확률분표가 오른쪽 표와 같을 때,

X	x_1	x_2	x_3	$\cdots$	x_n	합계
$\mathrm{P}(X=x)$	p_1	p_2	p_3	$\cdots$	p_n	1

$$x_1p_1 + x_2p_2 + x_3p_3 + \cdots + x_np_n = \sum_{i=1} x_ip_i$$

를 확률변수 X의 '기댓값' 또는 '평균'이라 하고, 기호로

$$\mathrm{E}(X)$$

와 같이 나타낸다.

☑ $\mathrm{E}(X)$의 E는 기댓값을 뜻하는 Expectation의 첫 글자이다.

4 이산확률변수 X의 기댓값(평균) ❷

예 한 개의 주사위를 두 번 던지는 시행에서 6의 약수의 눈이 나온 횟수를 확률변수 X라 할 때, X의 기댓값 $\mathrm{E}(X)$를 구해 보자. X가 갖는 값이 $0, 1, 2$이고 X의 확률질량함수는

$$\mathrm{P}(X=x) = {}_2\mathrm{C}_x \left(\frac{2}{3}\right)^x \left(\frac{1}{3}\right)^{2-x} \quad (x=0,\ 1,\ 2)$$

이다. 이때 확률변수 X의 확률분포를 표로 나타내면 오른쪽과 같다.

X	0	1	2	합계
$\mathrm{P}(X=x)$	$\frac{1}{9}$	$\frac{4}{9}$	$\frac{4}{9}$	1

따라서 $\mathrm{E}(X) = \frac{1}{9} \times (0\times1 + 1\times4 + 2\times4) = \frac{12}{9} = \frac{4}{3}$

5 이산확률변수 X의 분산, 표준편차 ❶

이산확률변수 X의 확률분포가 오른쪽 표와 같을 때,

X	x_1	x_2	x_3	$\cdots$	x_n	합계
$\mathrm{P}(X=x)$	p_1	p_2	p_3	$\cdots$	p_n	1

확률변수 X의 분산과 표준편차는 다음과 같다.

(1) 분산

$\mathrm{E}(X)=m$일 때, $(X-m)^2$의 평균

$$\mathrm{E}\left((X-m)^2\right)=\sum_{i=1}^{n}(x_i-m)^2p_i$$

$$=(x_1-m)^2p_1+(x_2-m)^2p_2+\cdots+(x_n-m)^2p_n$$

을 확률변수 X의 '분산'이라 하고, 기호로

5 이산확률변수 X의 분산, 표준편차 ❷

$$\mathrm{V}(X)$$

와 같이 나타낸다. 이때

$$\mathrm{V}(X)=\sum_{i=1}^{n}(x_i-m)^2p_i$$

$$=\sum_{i=1}^{n}\left({x_i}^2p_i-2mx_ip_i+m^2p_i\right)$$

$$=\sum_{i=1}^{n}{x_i}^2p_i-2m\sum_{i=1}^{n}x_ip_i+m^2\sum_{i=1}^{n}p_i$$

$$=\sum_{i=1}^{n}{x_i}^2p_i-2m^2+m^2\quad\left(\because\ \sum_{i=1}^{n}x_ip_i=m,\ \sum_{i=1}^{n}p_i=1\right)$$

5 이산확률변수 X의 분산, 표준편차 ❸

$$= \sum_{i=1}^{n} {x_i}^2 p_i - m^2$$

이므로 $\mathrm{V}(X) = \mathrm{E}(X^2) - \{\mathrm{E}(X)\}^2 = \mathrm{E}(X^2) - m^2$

(2) 표준편차

분산 $\mathrm{V}(X)$의 양의 제곱근을 확률변수 X의 '표준편차'라 하고, 기호로 $\sigma(X)$와 같이 나타낸다. 즉

$$\sigma(X) = \sqrt{\mathrm{V}(X)}$$

☑ $\mathrm{V}(X)$의 V는 분산을 뜻하는 Variance의 첫 글자이고 $\sigma(X)$의 σ는 표준편차를 뜻하는 standard deviation의 첫 글자 s에 해당하는 그리스 문자이다.

5 이산확률변수 X의 분산, 표준편차 ❹

예 확률변수 X의 확률분포가 오른쪽 표와 같을 때, X의 평균, 분산, 표준편차를 구해 보자.

X	2	3	4	합계
$\mathrm{P}(X=x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	1

$$m = \mathrm{E}(X) = \frac{1}{4} \times (2 \times 1 + 3 \times 2 + 4 \times 1) = \frac{12}{4} = 3$$

① $\mathrm{V}(X) = \mathrm{E}((X-m)^2) = \sum_{i=1}^{n} (x_i - m)^2 p_i$를 이용하면

$$\mathrm{V}(X) = \frac{1}{4} \times \{(-1)^2 + 0^2 + 1^2\} = \frac{2}{4} = \frac{1}{2}$$

5 이산확률변수 X의 분산, 표준편차 ❺

$$\sigma(X)=\sqrt{\mathrm{V}(X)}=\sqrt{\frac{1}{2}}=\frac{\sqrt{2}}{2}$$

② $\mathrm{V}(X)=\mathrm{E}\left(X^2\right)-\{\mathrm{E}(X)\}^2$을 이용하면

$$\mathrm{V}(X)=\frac{1}{4}\times\left(2^2\times1+3^2\times2+4^2\times1\right)-3^2$$

$$=\frac{19}{2}-9=\frac{1}{2}$$

$$\sigma(X)=\sqrt{\mathrm{V}(X)}=\sqrt{\frac{1}{2}}=\frac{\sqrt{2}}{2}$$

6 이산확률변수 $aX+b$의 평균, 분산, 표준편차 ❶

이산확률변수 X와 두 상수 a, b $(a\neq0)$에 대하여
이산확률변수 $aX+b$의 평균, 분산, 표준편차는 다음과 같다.

(1) $\mathrm{E}(aX+b)=a\mathrm{E}(X)+b$

(2) $\mathrm{V}(aX+b)=a^2\mathrm{V}(X)$

(3) $\sigma(aX+b)=|a|\sigma(X)$

☑ 이산확률변수 X의 확률분포가 다음과 같을 때,

X	x_1	x_2	x_3	$\cdots$	x_n	합계
$\mathrm{P}(X=x)$	p_1	p_2	p_3	$\cdots$	p_n	1

이산확률변수

$Y=aX+b$

(a와 b는 상수, $a\neq0$)

의 평균, 분산, 표준편차를 구해 보자.

6 이산확률변수 $aX+b$의 평균, 분산, 표준편차 ❷

$y_i = ax_i + b \ (i = 1, 2, 3, \cdots, n)$이라 할 때, 확률

$$\mathrm{P}(Y = y_i) = \mathrm{P}(X = x_i) = p_i$$

이므로 확률변수 Y의 확률분포를 표로 나타내면 오른쪽과 같다.

Y	y_1	y_2	y_3	$\cdots$	y_n	합계
$\mathrm{P}(Y=y)$	p_1	p_2	p_3	$\cdots$	p_n	1

따라서 확률변수 Y의 평균은

$$\mathrm{E}(Y) = \sum_{i=1}^{n} y_i p_i = \sum_{i=1}^{n} (ax_i + b) p_i$$

$$= a\sum_{i=1}^{n} x_i p_i + b\sum_{i=1}^{n} p_i = a\mathrm{E}(X) + b$$

확률변수 X의 평균을 m이라 하면 확률변수 Y의 평균은

6 이산확률변수 $aX+b$의 평균, 분산, 표준편차 ❸

$am+b$이므로 Y의 분산과 표준편차는

$$\mathrm{V}(Y) = \sum_{i=1}^{n} \{y_i - (am+b)\}^2 p_i$$

$$= \sum_{i=1}^{n} \{(ax_i + b) - (am+b)\}^2 p_i$$

$$= \sum_{i=1}^{n} a^2 (x_i - m)^2 p_i$$

$$= a^2 \sum_{i=1}^{n} (x_i - m)^2 p_i = a^2 \mathrm{V}(X)$$

$$\sigma(Y) = \sqrt{\mathrm{V}(Y)} = \sqrt{a^2 \mathrm{V}(X)} = |a|\sqrt{\mathrm{V}(X)} = |a|\sigma(X)$$

6 이산확률변수 $aX+b$의 평균, 분산, 표준편차 ❹

예 이산확률변수 X에 대하여 $\mathrm{E}(X)=2$, $\mathrm{V}(X)=9$일 때, 확률변수 $-2X+5$의 평균, 분산, 표준편차는 다음과 같다.

$$\mathrm{E}(-2X+5)=-2\mathrm{E}(X)+5=-2\times 2+5=1$$

$$\mathrm{V}(-2X+5)=(-2)^2\mathrm{V}(X)=4\times 9=36$$

$$\sigma(-2X+5)=|-2|\sigma(X)=2\times\sqrt{9}=6$$

7 이항분포 ❶

한 번의 시행에서 사건 A가 일어날 확률이 p로 일정할 때, n번의 독립시행에서 사건 A가 일어나는 횟수를 확률변수 X라 하면 X가 갖는 값은 $0, 1, 2, \cdots, n$이고, X의 확률질량함수는

$$\mathrm{P}(X=x)={}_n\mathrm{C}_x\,p^x q^{n-x}$$

$$(x=0,\ 1,\ 2,\ \cdots,\ n\text{이고 } q=1-p)$$

이다. 이와 같은 이산확률변수 X의 확률분포를 ‘이항분포’라 하고, 기호로 $\mathrm{B}(n,\ p)$와 같이 나타낸다. 이때 확률변수 X는 ‘이항분포 $\mathrm{B}(n,\ p)$를 따른다’고 하며, X의 확률분포를 표로 나타내면 다음과 같다.

X	0	1	2	$\cdots$	x	$\cdots$	n	합계
$\mathrm{P}(X=x)$	${}_n\mathrm{C}_0 p^0 q^n$	${}_n\mathrm{C}_1 p^1 q^{n-1}$	${}_n\mathrm{C}_2 p^2 q^{n-2}$	$\cdots$	${}_n\mathrm{C}_x p^x q^{n-x}$	$\cdots$	${}_n\mathrm{C}_n p^n q^0$	1

7 이항분포 ❷

☑ (1) 위의 표에서 각 확률은 이항정리에 의하여 $(p+q)^n$을 전개한 식

$$(p+q)^n = {}_n\mathrm{C}_0\, p^0 q^n + {}_n\mathrm{C}_1\, p^1 q^{n-1} + {}_n\mathrm{C}_2\, p^2 q^{n-2} + \cdots + {}_n\mathrm{C}_x\, p^x q^{n-x} + \cdots + {}_n\mathrm{C}_n\, p^n q^0$$

의 우변의 각 항과 같다. 이때 $p+q=1$이므로

$\displaystyle\sum_{x=0}^{n} {}_n\mathrm{C}_x\, p^x q^{n-x} = 1$임을 알 수 있다.

(2) 이항분포 $\mathrm{B}(n,\ p)$의 B는 이항분포를 뜻하는 Binomial distribution의 첫 글자이다.

7 이항분포 ❸

예 한 개의 주사위를 한 번 던지는 시행에서 6의 약수의 눈이 나오는 사건을 A라 하고, 이 주사위를 36번 던질 때 사건 A가 일어나는 횟수를 확률변수 X라 하자. 주사위를 한 번 던질 때 사건 A가 일어날 확률이 $\frac{2}{3}$이고 독립시행의 횟수가 36이므로 확률변수 X는 이항분포 $\mathrm{B}\left(36,\ \frac{2}{3}\right)$를 따른다.

$\boxed{8}$ 이항분포의 평균, 분산, 표준편차

확률변수 X가 이항분포 $\mathrm{B}(n,\ p)$를 따를 때,

(1) 평균 : $\mathrm{E}(X) = np$

(2) 분산 : $\mathrm{V}(X) = npq$ (단, $q = 1 - p$)

(3) 표준편차 : $\sigma(X) = \sqrt{\mathrm{V}(X)} = \sqrt{npq}$ (단, $q = 1 - p$)

✡ 이항분포의 평균, 분산, 표준편차

$$\mathrm{P}(X = x) = {}_n\mathrm{C}_x\, p^x (1-p)^{n-x} \quad (x = 0,\ 1,\ 2,\ \cdots,\ n)$$

(o) $\displaystyle\sum_{x=0}^{n} {}_n\mathrm{C}_x\, p^x (1-p)^{n-x} = (p + 1 - p)^n = 1$

(1) $\displaystyle\mathrm{E}(X) = \sum_{x=0}^{n} x\, {}_n\mathrm{C}_x\, p^x (1-p)^{n-x}$

$$= \sum_{x=1}^{n} \frac{n!}{(x-1)!\,(n-x)!} p^x (1-p)^{n-x} \Rightarrow x - 1 \overset{\text{def}}{=} l$$

$$= np \sum_{l=0}^{n-1} \frac{(n-1)!}{l!\,(n-1-l)!} p^l (1-p)^{n-1-l}$$

$$= np\{p + (1-p)\}^{n-1} = np$$

(2) $\displaystyle\mathrm{E}\{X(X-1)\} = \sum_{x=0}^{n} x(x-1)\, {}_n\mathrm{C}_x\, p^x (1-p)^{n-x}$

$$= \sum_{x=2}^{n} \frac{n!}{(x-2)!\,(n-x)!} p^x (1-p)^{n-x} \Rightarrow x - 2 \overset{\text{def}}{=} l$$

$$= n(n-1)p^2 \sum_{l=0}^{n-2} \frac{(n-2)!}{l!\,(n-2-l)!} p^l (1-p)^{n-2-l}$$

$$= n(n-1)p^2 \{p + (1-p)\}^{n-2} = n(n-1)p^2$$

$$\mathrm{E}\left(X^2\right) = \mathrm{E}\{X(X-1)\} + \mathrm{E}(X)$$

$$= n(n-1)p^2 + np = n^2p^2 + np(1-p)$$

(3) $\mathrm{E}\left(X^2\right) = n^2p^2 + np(1-p)$

$\therefore\ \mathrm{V}(X) = \mathrm{E}\left(X^2\right) - \{\mathrm{E}(X)\}^2$

$$= n^2p^2 + np(1-p) - (np)^2$$

$$= npq \quad (\text{단, } q = 1-p)$$

(4) $\sigma(X) = \sqrt{npq}$ (단, $q = 1-p$)

$\boxed{9}$ 큰수의 법칙

어떤 시행에서 사건 A가 일어날 수학적 확률이 p일 때, n번의 독립시행에서 사건 A가 일어나는 횟수를 확률변수 X라 하면 임의의 양수 h에 대하여 n의 값이 한없이 커질 때, 확률 $\mathrm{P}\left(\left|\frac{X}{n} - p\right| < h\right)$의 값은 1에 한없이 가까워진다.

☑ 큰수의 법칙에 의하여 시행 횟수 n이 충분히 클 때, 사건 A의 상대도수는 수학적 확률에 가까워지므로 사건 A의 상대도수 $\frac{X}{n}$를 사건 A가 일어날 확률 $\mathrm{P}(A)$로 간주할 수 있다. 따라서 자연 현상이나 사회 현상에서 수학적 확률을 구하기 어려운 경우에는 시행 횟수를 충분히 크게 한 후 사건의 상대도수를 구하여 수학적 확률로 이용할 수 있다.

✡ 이항분포의 평균과 분산 ❶

확률변수 X가 이항분포 $\mathrm{B}(n,\ p)$를 따를 때, X의 평균 $\mathrm{E}(X)$와 분산 $\mathrm{V}(X)$를 미분을 이용하여 구할 수 있다.

(단, $p+q=1$)

(1) 이항정리에 의하여 $(x+q)^n$을 다음과 같이 나타낼 수 있다.

$$(x+q)^n=\sum_{r=0}^{n} {}_{n}\mathrm{C}_{r}\, x^r q^{n-r}$$

양변을 x에 대하여 미분하면

$$n(x+q)^{n-1}=\sum_{r=0}^{n} r\,{}_{n}\mathrm{C}_{r}\, x^{r-1} q^{n-r}$$

양변에 x를 곱하면

✡ 이항분포의 평균과 분산 ❷

$$nx(x+q)^{n-1}=\sum_{r=0}^{n} r\,{}_{n}\mathrm{C}_{r}\, x^{r} q^{n-r} \qquad \cdots\cdots ㉠$$

㉠에 $x=p$를 대입하면

$$np(p+q)^{n-1}=\sum_{r=0}^{n} r\,{}_{n}\mathrm{C}_{r}\, p^{r} q^{n-r}$$

이때 $\mathrm{E}(X)=\sum_{r=0}^{n} r\,{}_{n}\mathrm{C}_{r}\, p^{r} q^{n-r}$

$$\mathrm{E}(X)=np$$

(2) $\mathrm{V}(X)=\mathrm{E}\left(X^2\right)-\{\mathrm{E}(X)\}^2$이므로

㉠의 양변을 x에 대하여 미분하면

✡ 이항분포의 평균과 분산 ❸

$$n(x+q)^{n-1}+n(n-1)x(x+q)^{n-2}$$

$$=\sum_{r=0}^{n} r^2 {}_{n}\mathrm{C}_{r}\, x^{r-1}q^{n-r}$$

양변에 x를 곱하면

$$nx(x+q)^{n-1}+n(n-1)x^2(x+q)^{n-2}$$

$$=\sum_{r=0}^{n} r^2 {}_{n}\mathrm{C}_{r}\, x^{r}q^{n-r} \qquad \cdots\cdots ㉡$$

㉡에 $x=p$를 대입하면

$$np(p+q)^{n-1}+n(n-1)p^2(p+q)^{n-2}$$

$$=\sum_{r=0}^{n} r^2 {}_{n}\mathrm{C}_{r}\, p^{r}q^{n-r}$$

✡ 이항분포의 평균과 분산 ❹

$$np+n^2p^2-np^2=\sum_{r=0}^{n} r^2 {}_{n}\mathrm{C}_{r}\, p^{r}q^{n-r}$$

이때 $\mathrm{E}\left(X^2\right)=\sum_{r=0}^{n} r^2 {}_{n}\mathrm{C}_{r}\, p^{r}q^{n-r}$이므로

$$\mathrm{E}\left(X^2\right)=np+n^2p^2-np^2$$

따라서 분산 $\mathrm{V}(X)$는

$$\begin{aligned}\mathrm{V}(X)&=\mathrm{E}\left(X^2\right)-\{\mathrm{E}(X)\}^2\\&=np+n^2p^2-np^2-(np)^2\\&=np(1-p)=npq\end{aligned}$$