수지, 이산확률분포

2018년 10월 17일

예시 1) 중학교 통계 복습

다섯 개의 숫자 1, 2, 3, 4, 5의 평균 m은

$$m = \frac{1+2+3+4+5}{5} = 3$$

이다. 분산 V는 '편차의 제곱의 평균'이다. 이때, 편차란 '변량 − 평균'이다. 따라서

변량	1	2	3	4	5
편차	−2	−1	3	1	2
편차^2	4	1	0	1	4

이고

$$V = \frac{4+1+0+1+4}{5} = \boxed{(1)}$$

표준편차 σ는 $\sigma = \sqrt{V} = \boxed{(2)}$ 이다.

예시 2) 확률질량함수

검은 바둑돌이 2개, 흰 바둑돌이 2개 들어있는 주머니에서 2개의 바둑돌을 동시에 꺼낼 때, 꺼낸 흰 바둑돌의 개수를 X라고 하자. 그러면

$$P(X=0) = \frac{{}_2C_2 \times {}_2C_0}{{}_4C_2} = \frac{1}{6}$$
$$P(X=1) = \boxed{(3)}$$
$$P(X=2) = \boxed{(4)}$$

이다. 이 함수 $P(X=x)$를 **확률질량함수**라고 부르고 다음과 같이 표로 표현한다.

X	0	1	2	합계
$P(X=x)$	$\frac{1}{6}$	$\boxed{(3)}$	$\boxed{(4)}$	1

이때 X의 평균 $E(X)$는

$$E(X) = 0 \times \frac{1}{6} + 1 \times \boxed{(3)} + 2 \times \boxed{(4)} = 1$$

이고, X의 분산 $V(X)$는 편차 $X-1$의 제곱의 평균이므로

$(X-1)^2$	1	0	1	합계
$P(X=x)$	$\frac{1}{6}$	$\boxed{(3)}$	$\boxed{(4)}$	1

에서

$$V(X) = 1 \times \frac{1}{6} + 0 \times \boxed{(3)} + 1 \times \boxed{(4)} = \boxed{(5)}$$

이다. 또 표준편차 $\sigma(X)$는

$$\sigma(X) = \sqrt{V(X)} = \frac{\sqrt{3}}{3}.$$

정의 3)
확률변수 X의 확률질량함수가

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

로 주어질 때, X의 평균 $E(X)$는

$$\begin{aligned} E(X) &= x_1p_1 + x_2p_2 + \cdots + x_np_n \\ &= \sum_{i=1}^{n} x_ip_i \end{aligned}$$

이다. $E(X) = m$이라고 하면, X의 분산 $V(X)$는 편차 $X - m$의 제곱의 평균이므로

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

이다. X의 표준편차 $\sigma(X)$는

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

이다.

정리 4)

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

증명)

$$\begin{aligned} V(X) &= \sum_{i=1}^{n}(x_i - m)^2 p_i \\ &= \sum_{i=1}^{n}({x_i}^2 - 2mx_i + m^2)p_i \\ &= \sum_{i=1}^{n}({x_i}^2 p_i - 2mx_i p_i + m^2 p_i) \\ &= \sum_{i=1}^{n}{x_i}^2 p_i - 2m\sum_{i=1}^{n}x_i p_i + m^2\sum_{i=1}^{n}p_i \end{aligned}$$

이때,

$$\sum_{i=1}^{n}{x_i}^2 p_i = \boxed{(6)}, \quad \sum_{i=1}^{n}x_i p_i = \boxed{(7)}, \qquad \sum_{i=1}^{n}p_i = \boxed{(8)}$$

이므로

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

□

예시 5)

예시 2)의 분산을 정리 4)의 방법으로 구하자. 먼저 $E(X^2)$를 구하면,

X^2	0	1	4	합계
$P(X^2 = x^2)$	$\frac{1}{6}$	$\frac{2}{3}$	$\frac{1}{6}$	1

에서

$$E(X^2) = 0 \times \frac{1}{6} + 1 \times \frac{2}{3} + 4 \times \frac{1}{6} = \frac{4}{3}$$

따라서

$$V(X) = E(X^2) - \{E(X)\}^2 = \frac{4}{3} - 1^2 = \frac{1}{3}$$

이다. 이것은 예시 2)의 결과와도 일치한다.

정리 6)
$Y = aX + b$일 때,

$$E(Y) = aE(X) + b, \quad V(Y) = a^2V(X)$$

이다.

증명)
$X = x_i$이면 $Y = ax_i + b$이다. 따라서 Y의 확률질량함수는

Y	$ax_1 + b$	$ax_2 + b$	$ax_3 + b$	$\cdots$	$ax_n + b$	합계
$P(Y = y)$	p_1	p_2	p_3	$\cdots$	p_n	1

따라서

$$\begin{aligned} E(Y) &= \sum_{i=1}^{n}(ax_i + b)p_i \\ &= \sum_{i=1}^{n}(ax_ip_i + bp_i) \\ &= a\sum_{i=1}^{n}x_ip_i + b\sum_{i=1}^{n}p_i \end{aligned}$$

이때,

$$\sum_{i=1}^{n}x_ip_i = \boxed{(7)}, \qquad \sum_{i=1}^{n}p_i = \boxed{(8)}$$

이므로

$$E(Y) = aE(X) + b$$

이다.

또

$$V(Y) = E(Y^2) - \{E(Y)\}^2$$

에서

$$\begin{aligned} E(Y^2) &= \sum_{i=1}^{n} (ax_i + b)^2 p_i \\ &= \sum_{i=1}^{n} (a^2 {x_i}^2 + 2abx_i + b^2) p_i \\ &= a^2 \sum_{i=1}^{n} {x_i}^2 p_i + 2ab \sum_{i=1}^{n} x_i p_i + b^2 \sum_{i=1}^{n} p_i \\ &= \boxed{\quad (9) \quad} \end{aligned}$$

이고

$$\begin{aligned} \{E(Y)\}^2 &= \{aE(X) + b\}^2 \\ &= (am + b)^2 \\ &= \boxed{\quad (10) \quad} \end{aligned}$$

이므로

$$\begin{aligned} V(Y) &= E(Y^2) - \{E(Y)\}^2 \\ &= \boxed{\quad (9) \quad} - \boxed{\quad (10) \quad} \\ &= a^2 E(X^2) - a^2 m^2 \\ &= a^2 \{E(X^2) - m^2\} \\ &= a^2 \left[E(X^2) - \{E(X)\}^2 \right] \\ &= a^2 V(X) \end{aligned}$$

□

예시 7)

주사위를 90번 던져서 1의 눈이 나온 횟수를 X라고 하자. 그러면

$$
\begin{aligned}
P(X=0) &= \left(\frac{5}{6}\right)^{90} \\
P(X=1) &=_{90} C_1 \left(\frac{1}{6}\right)^{1} \left(\frac{5}{6}\right)^{89} \\
P(X=2) &=_{90} C_2 \left(\frac{1}{6}\right)^{2} \left(\frac{5}{6}\right)^{88} \\
&\vdots \\
P(X=89) &=_{90} C_{89} \left(\frac{1}{6}\right)^{89} \left(\frac{5}{6}\right)^{1} \\
P(X=90) &= \left(\frac{1}{6}\right)^{90}
\end{aligned}
$$

즉

$$P(X=x) =_{90} C_x \left(\frac{1}{6}\right)^{x} \left(\frac{5}{6}\right)^{90-x}$$

이다.

한 번 던질 때마다 1의 눈이 나올 확률은 $\frac{1}{6}$이므로, 1의 눈이 나오는 횟수는 평균적으로 $90 \times \frac{1}{6} = 15$번 임을 예상할 수 있다. 즉

$$E(X) = 15$$

일 것이다.

정의 8)

일어날 확률이 p인 사건의 시행을 n회 반복할 때, 이 사건이 일어난 횟수를 X라고 하자. X의 확률질량함수는

$$P(X=x) =_n C_x p^x q^{n-x}$$

이다. 이와 같은 X의 확률분포를 이항분포라고 하며 기호로

$$X \sim B(n,p)$$

라고 쓴다.

정리 9)

$X \sim B(n,p)$이면

$$E(X) = np, \quad V(X) = npq$$

이다. $(q = 1-p)$

이 식의 증명에는 $r \cdot {}_nC_r = n \cdot {}_{n-1}C_{r-1}$이 필요하다. 그리고 이 식은 쉽게 증명된다 ;

$$\begin{aligned} r \cdot {}_nC_r &= r \cdot \frac{n!}{r!(n-r)!} = \frac{n!}{(r-1)!(n-r)!} \\ &= n \cdot \frac{\boxed{(11)}}{(r-1)!\{(n-1)-(r-1)\}!} \\ &= n \cdot {}_{n-1}C_{r-1} \end{aligned}$$

증명)

X의 확률질량함수는

X	0	1	2	$\cdots$	n	합계
$P(X=x)$	p_0	p_1	p_2	$\cdots$	p_n	1

이다. 이때,

$$p_r = P(X=r) = {}_nC_r p^r q^{n-r}$$

이다. 그러면

$$\begin{aligned} E(X) &= \sum_{r=0}^{n} r \cdot p_r = \sum_{r=0}^{n} r \cdot {}_nC_r p^r q^{n-r} \\ &= \boxed{(12)} + \sum_{r=1}^{n} r \cdot {}_nC_r p^r q^{n-r} \\ &= \sum_{r=1}^{n} n \cdot {}_{n-1}C_{r-1} p^r q^{n-r} \\ &= np \sum_{r=1}^{n} {}_{n-1}C_{r-1} p^{r-1} q^{(n-1)-(r-1)} \end{aligned}$$

$r-1 = s$로 치환하면 $s = r+1$이므로

$$\begin{aligned} E(X) &= np \sum_{s=0}^{n-1} {}_{n-1}C_s p^s q^{(n-1)-s} \\ &= np(p+q)^{n-1} = \boxed{(13)}. \end{aligned}$$

한편,

$$\begin{aligned}
E(X^2) &= \sum_{r=0}^{n} r^2 \cdot p_r \\
&= \sum_{r=0}^{n} r^2 \cdot_n C_r p^r q^{n-r} \\
&= \sum_{r=0}^{n} \{r(r-1)+r\} \cdot_n C_r p^r q^{n-r} \\
&= \sum_{r=0}^{n} r(r-1) \cdot_n C_r p^r q^{n-r} + \sum_{r=0}^{n} r \cdot_n C_r p^r q^{n-r} \\
&= \boxed{(12)} + \boxed{(12)} + \sum_{r=2}^{n} r(r-1) \cdot_n C_r p^r q^{n-r} + \boxed{(13)}
\end{aligned}$$

이때,

$$r(r-1) \cdot_n C_r = (r-1) \times n \cdot_{n-1} C_{r-1} = n(n-1) \cdot_{n-2} C_{r-2}$$

이므로

$$\begin{aligned}
E(X^2) &= \sum_{r=2}^{n} n(n-1) \cdot_{n-2} C_{r-2} p^r q^{n-r} + \boxed{(13)} \\
&= n(n-1)p^2 \sum_{r=2}^{n} \cdot_{n-2} C_{r-2} p^{r-2} q^{(n-2)-(r-2)} + \boxed{(13)}
\end{aligned}$$

$r-2=s$로 치환하면 $s=r+2$이므로

$$\begin{aligned}
E(X^2) &= n(n-1)p^2 \sum_{s=0}^{n-2} {}_{n-2}C_s p^s q^{(n-2)-s} + \boxed{(13)} \\
&= n(n-1)p^2 (p+q)^{n-2} + \boxed{(13)} \\
&= \boxed{(14)} + \boxed{(13)}.
\end{aligned}$$

이다.

따라서

$$\begin{aligned}
V(X) &= E(X^2) - \{E(X)\}^2 \\
&= \left(\boxed{(14)} + \boxed{(13)}\right) - \boxed{(13)}^2 \\
&= npq.
\end{aligned}$$

답

(1) 2

(2) $\sqrt{2}$

(3) $\frac{2}{3}$

(4) $\frac{1}{6}$

(5) $\frac{1}{3}$

(6) $E(X^2)$

(7) $E(X$ 또는 m

(8) 1

(9) $a^2E(X^2) + 2abm + b^2$

(10) $a^2m^2 + 2abm + b^2$

(11) $(n-1)!$

(12) 0

(13) np

(14) $n(n-1)p^2$