# 01 여러 가지 순열

## 1. 원순열

(1) 원순열의 뜻

서로 다른 것을 원형으로 배열하는 순열을 원순열이라고 한다.

(2) 원순열의 수

서로 다른 $n$개를 원형으로 배열하는 원순열의 수는

$$\frac{n!}{n}=(n-1)!$$

**설명** 원순열에서는 회전하여 일치하는 것은 모두 같은 것으로 본다.

서로 다른 $n$개를 일렬로 나열하는 순열의 수는 $n!$이고, 이를 원형으로 배열하면 같은 것이 $n$가지씩 있으므로 서로 다른 $n$개를 원형으로 배열하는 원순열의 수는

$$\frac{n!}{n}=(n-1)!$$

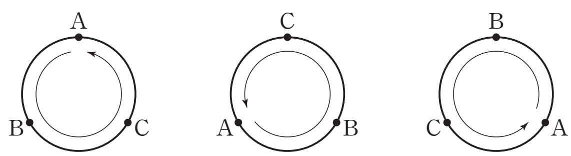
**예** 세 개의 문자 A, B, C를 원형으로 배열하는 경우의 수를 구해 보자.

세 개의 문자 A, B, C를 일렬로 나열하면

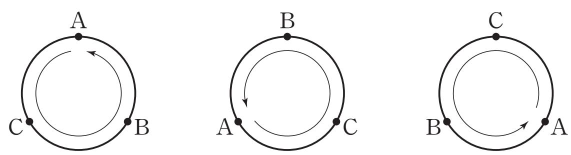
ABC, ACB, BAC, BCA, CAB, CBA

이므로 그 경우의 수는 $3!$

이때 ABC, CAB, BCA를 원형으로 배열한 후 회전하면 서로 일치하므로 같은 경우이다.




또 ACB, BAC, CBA를 원형으로 배열한 후 회전하면 서로 일치하므로 같은 경우이다.




따라서 세 개의 문자 A, B, C를 일렬로 나열하는 경우의 수는 $3!$이지만 이를 원형으로 배열하면 회전하여 같아지는 것이 3가지씩 있으므로 세 개의 문자 A, B, C를 원형으로 배열하는 원순열의 수는

$$\frac{3!}{3}=(3-1)!=2!=2$$

(3) 서로 다른 $n$개에서 $r$ $(0<r\le n)$개를 택하여 원형으로 배열하는 경우의 수는

$$\frac{{}_{n}\mathrm{P}_{r}}{r}$$

**설명** 서로 다른 $n$개에서 $r$개를 택하는 조합의 수는 ${}_{n}\mathrm{C}_{r}$이고,

서로 다른 $r$개를 원형으로 배열하는 원순열의 수는 $(r-1)!$이므로

서로 다른 $n$개에서 $r$개를 택하여 원형으로 배열하는 경우의 수는

$${}_{n}\mathrm{C}_{r}\times(r-1)!={}_{n}\mathrm{C}_{r}\times\frac{r!}{r}=\frac{{}_{n}\mathrm{C}_{r}\times r!}{r}=\frac{{}_{n}\mathrm{P}_{r}}{r}$$

## 01 여러 가지 순열

### 2. 중복순열

(1) 중복순열의 뜻

서로 다른 $n$개에서 중복을 허락하여 $r$개를 택하여 일렬로 나열하는 것을 $n$개에서 $r$개를 택하는 중복순열이라 하고, 이 중복순열의 수를 기호로

$${}_n\Pi_r$$

와 같이 나타낸다.

**참고** 기호 ${}_n\Pi_r$에서 $\Pi$는 곱을 뜻하는 영어 Product의 첫 글자인 P에 해당하는 그리스 문자로 ‘파이’라고 읽는다.

(2) 중복순열의 수

서로 다른 $n$개에서 $r$개를 택하는 중복순열의 수는

$${}_n\Pi_r=n^r$$

**설명** 서로 다른 $n$개에서 중복을 허락하여 $r$개를 택하여 일렬로 나열할 때, 첫 번째, 두 번째, 세 번째, $\cdots$, $r$ 번째 자리에 올 수 있는 것은 각각 $n$가지씩이다.

| 나열 순서 | 첫 번째 | 두 번째 | 세 번째 | $\cdots$ | $r$ 번째 |
|---|---|---|---|---|---|
| | ↓ | ↓ | ↓ | | ↓ |
| 경우의 수 | $n$ | $n$ | $n$ | $\cdots$ | $n$ |

따라서 중복순열의 수 ${}_n\Pi_r$는 곱의 법칙에 의하여

$${}_n\Pi_r=\underbrace{n\times n\times n\times\cdots\times n}_{r\text{개}}=n^r$$

**예1** ${}_3\Pi_4=3^4=81$, ${}_4\Pi_3=4^3=64$

**예2** 숫자 1, 2, 3, 4, 5 중에서 중복을 허락하여 3개의 숫자를 택해 일렬로 나열하여 만들 수 있는 세 자리 자연수의 개수를 구해 보자.

백의 자리, 십의 자리, 일의 자리에 올 수 있는 숫자는 각각 1, 2, 3, 4, 5의 5가지이다.

| 백의 자리 | 십의 자리 | 일의 자리 |
|---|---|---|
| 1, 2, 3, 4, 5 | 1, 2, 3, 4, 5 | 1, 2, 3, 4, 5 |
| ↓ | ↓ | ↓ |
| 5가지 | 5가지 | 5가지 |

따라서 만들 수 있는 세 자리 자연수의 개수는 곱의 법칙에 의하여

$$5\times5\times5=125$$

이고, 이것은 서로 다른 5개에서 중복을 허락하여 3개를 택해 일렬로 나열하는 중복순열의 수와 같다.

즉, ${}_5\Pi_3=5^3=125$이다.

**참고** 공집합이 아닌 두 집합 $X$, $Y$의 원소의 개수가 각각 $m$, $n$일 때, 집합 $X$에서 집합 $Y$로의 모든 함수의 개수는 공역 $Y$의 서로 다른 $n$개의 원소 중에서 중복을 허락하여 $m$개를 택하여 일렬로 나열하는 중복순열의 수와 같으므로

$${}_n\Pi_m=n^m$$

# 01 여러 가지 순열

## 3. 같은 것이 있는 순열

(1) 같은 것이 있는 순열의 뜻

같은 것이 포함되어 있는 $n$개를 일렬로 나열하는 것을 같은 것이 있는 순열이라고 한다.

(2) 같은 것이 있는 순열의 수

$n$개 중에서 같은 것이 각각 $p$개, $q$개, $\cdots$, $r$개씩 있을 때, $n$개를 일렬로 나열하는 순열의 수는

$$\frac{n!}{p!\times q!\times\cdots\times r!}\ (\text{단, } p+q+\cdots+r=n)$$

**설명** 5개의 문자 $a$, $a$, $a$, $b$, $b$를 일렬로 나열하는 경우의 수를 구해 보자.

5개의 문자 $a$, $a$, $a$, $b$, $b$에서 3개의 $a$를 구별하여 각각 $a_1$, $a_2$, $a_3$이라 하고, 2개의 $b$를 구별하여 각각 $b_1$, $b_2$라 하면 5개의 문자 $a_1$, $a_2$, $a_3$, $b_1$, $b_2$를 일렬로 나열하는 경우의 수는

$${}_5\mathrm{P}_5=5!$$

그런데 $5!$가지 중에서 다음과 같은 $3!\times2!$가지의 서로 다른 순열은 번호를 이용한 구별이 없다면 모두 $aaabb$와 같다.

| | |
|---|---|
| $a_1a_2a_3b_1b_2$ | $a_1a_2a_3b_2b_1$ |
| $a_1a_3a_2b_1b_2$ | $a_1a_3a_2b_2b_1$ |
| $a_2a_1a_3b_1b_2$ | $a_2a_1a_3b_2b_1$ |
| $a_2a_3a_1b_1b_2$ | $a_2a_3a_1b_2b_1$ |
| $a_3a_1a_2b_1b_2$ | $a_3a_1a_2b_2b_1$ |
| $a_3a_2a_1b_1b_2$ | $a_3a_2a_1b_2b_1$ |

➡ $aaabb$

이와 같이 생각하면 5개의 문자 $a$, $a$, $a$, $b$, $b$를 일렬로 나열하는 순열의 수는

$$\frac{5!}{3!\times2!}=10$$

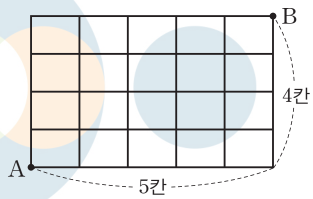
**예** 여섯 개의 숫자 1, 2, 2, 3, 3, 3을 일렬로 나열하여 만들 수 있는 여섯 자리 자연수의 개수는

$$\frac{6!}{1!\times2!\times3!}=60$$

**참고** 직사각형 모양으로 연결된 도로망을 따라 두 지점 사이를 최단 거리로 이동하는 경우의 수는 가로 방향으로 한 칸 움직이는 이동과 세로 방향으로 한 칸 움직이는 이동을 필요한 횟수만큼 일렬로 나열하는 경우의 수와 같으므로 같은 것이 있는 순열의 수를 이용하여 구할 수 있다.

오른쪽 그림과 같이 직사각형 모양으로 연결된 도로망에서 가로 방향의 칸의 수가 5, 세로 방향의 칸의 수가 4일 때, 이 도로망을 따라 A지점을 출발하여 B지점까지 최단 거리로 이동하는 경우의 수는

$$\frac{(5+4)!}{5!\times4!}=126$$



# 02 중복조합과 이항정리

## 1. 중복조합

(1) 중복조합의 뜻

서로 다른 $n$개에서 중복을 허락하여 $r$개를 택하는 조합을 중복조합이라 하고, 이 중복조합의 수를 기호로

$${}_{n}\mathrm{H}_{r}$$

와 같이 나타낸다.

**참고** ${}_{n}\mathrm{H}_{r}$에서 H는 Homogeneous의 첫 글자이다.

(2) 중복조합의 수

서로 다른 $n$개에서 $r$개를 택하는 중복조합의 수는

$${}_{n}\mathrm{H}_{r}={}_{n+r-1}\mathrm{C}_{r}$$

**설명** 세 개의 문자 $a$, $b$, $c$에서 중복을 허락하여 2개를 택하는 조합은

$$aa,\ ab,\ ac,\ bb,\ bc,\ cc$$

의 6가지이므로 ${}_{3}\mathrm{H}_{2}=6$이다.

이때 위의 6가지 경우를 문자가 들어갈 두 개의 자리 ○와 서로 다른 문자 사이를 구분할 두 개의 막대 ▮를 이용하여 오른쪽 그림과 같이 나타낼 수 있다.

즉, ○○▮▮를 일렬로 나열한 후 왼쪽에 놓인 ▮의 왼쪽에 ○가 있으면 그 자리에 문자 $a$를, ▮와 ▮ 사이에 ○가 있으면 그 자리에 문자 $b$를, 오른쪽에 놓인 ▮의 오른쪽에 ○가 있으면 그 자리에 문자 $c$를 넣으면 된다.

○○▮▮ ➡ $aa$
○▮○▮ ➡ $ab$
○▮▮○ ➡ $ac$
▮○○▮ ➡ $bb$
▮○▮○ ➡ $bc$
▮▮○○ ➡ $cc$

따라서 서로 다른 3개에서 2개를 택하는 중복조합의 수 ${}_{3}\mathrm{H}_{2}$는 2개의 ○와 2개의 ▮를 일렬로 나열하는 경우의 수와 같으므로 같은 것이 있는 순열의 수에 의하여

$${}_{3}\mathrm{H}_{2}=\frac{4!}{2!\times 2!}=\frac{4!}{2!\times(4-2)!}={}_{4}\mathrm{C}_{2}$$

이다.

일반적으로 서로 다른 $n$개에서 중복을 허락하여 $r$개를 택하는 중복조합의 수는 $r$개의 자리 ○와 $n$개를 구분하는 $(n-1)$개의 막대 ▮를 일렬로 나열하는 경우의 수와 같으므로

$${}_{n}\mathrm{H}_{r}=\frac{\{r+(n-1)\}!}{r!\times(n-1)!}={}_{r+(n-1)}\mathrm{C}_{r}={}_{n+r-1}\mathrm{C}_{r}$$

이다.

**참고** 서로 다른 $n$개에서 중복을 허락하여 $r$개를 택하는 중복조합의 수는 $r$개의 자리 ○와 $(n-1)$개의 막대 ▮를 놓을 $r+(n-1)=n+r-1$(개)의 자리 중에서 ○를 놓을 $r$개의 자리를 택하는 조합의 수와 같으므로

$${}_{n}\mathrm{H}_{r}={}_{n+r-1}\mathrm{C}_{r}$$

이다.

**예** 숫자 1, 2, 3에서 중복을 허락하여 5개를 택하는 경우의 수는 서로 다른 3개에서 5개를 택하는 중복조합의 수와 같으므로

$${}_{3}\mathrm{H}_{5}={}_{3+5-1}\mathrm{C}_{5}={}_{7}\mathrm{C}_{5}={}_{7}\mathrm{C}_{2}=\frac{7\times 6}{2\times 1}=21$$

이다.

## 2. 중복조합의 활용

(1) 방정식을 만족시키는 음이 아닌 정수해의 개수

방정식 $x_1+x_2+x_3+\cdots+x_n=r$ ($n$은 자연수, $r$는 음이 아닌 정수)를 만족시키는 음이 아닌 정수 $x_1$, $x_2$, $x_3$, $\cdots$, $x_n$의 모든 순서쌍 $(x_1, x_2, x_3, \cdots, x_n)$의 개수는

$${}_n\mathrm{H}_r$$

이다.

**설명** 방정식 $x+y+z+w=10$을 만족시키는 음이 아닌 정수 $x$, $y$, $z$, $w$의 모든 순서쌍 $(x, y, z, w)$의 개수를 구해 보자.

예를 들어 방정식 $x+y+z+w=10$의 해 중 하나인 $x=1$, $y=2$, $z=3$, $w=4$는 서로 다른 4개의 문자 $x$, $y$, $z$, $w$ 중에서 $x$를 1개, $y$를 2개, $z$를 3개, $w$를 4개 택한 것으로 생각할 수 있다.

같은 방법으로 생각하면 방정식 $x+y+z+w=10$의 모든 해의 순서쌍 $(x, y, z, w)$의 개수는 4개의 문자 $x$, $y$, $z$, $w$ 중에서 10개를 택하는 중복조합의 수와 같으므로 구하는 모든 순서쌍 $(x, y, z, w)$의 개수는

$${}_4\mathrm{H}_{10}={}_{4+10-1}\mathrm{C}_{10}={}_{13}\mathrm{C}_{10}={}_{13}\mathrm{C}_3=\frac{13\times12\times11}{3\times2\times1}=286$$

일반적으로 방정식 $x_1+x_2+x_3+\cdots+x_n=r$ ($n$은 자연수, $r$는 음이 아닌 정수)를 만족시키는 음이 아닌 정수 $x_1$, $x_2$, $x_3$, $\cdots$, $x_n$의 모든 순서쌍 $(x_1, x_2, x_3, \cdots, x_n)$의 개수는 서로 다른 $n$개의 문자 $x_1$, $x_2$, $x_3$, $\cdots$, $x_n$ 중에서 $r$개를 택하는 중복조합의 수 ${}_n\mathrm{H}_r$와 같다.

**참고** 방정식 $x_1+x_2+x_3+\cdots+x_n=r$ ($n$은 자연수, $r$는 $n$ 이상의 자연수)를 만족시키는 자연수 $x_1$, $x_2$, $x_3$, $\cdots$, $x_n$의 모든 순서쌍 $(x_1, x_2, x_3, \cdots, x_n)$의 개수는 다음과 같이 구한다.

$$x_1=x_1'+1,\ x_2=x_2'+1,\ x_3=x_3'+1,\ \cdots,\ x_n=x_n'+1$$

로 놓으면 $x_1'$, $x_2'$, $x_3'$, $\cdots$, $x_n'$은 음이 아닌 정수이고, 방정식 $x_1+x_2+x_3+\cdots+x_n=r$를 만족시키는 자연수 $x_1$, $x_2$, $x_3$, $\cdots$, $x_n$의 모든 순서쌍 $(x_1, x_2, x_3, \cdots, x_n)$의 개수는

방정식 $x_1'+x_2'+x_3'+\cdots+x_n'=r-n$을 만족시키는 음이 아닌 정수 $x_1'$, $x_2'$, $x_3'$, $\cdots$, $x_n'$의 모든 순서쌍 $(x_1', x_2', x_3', \cdots, x_n')$의 개수와 같으므로

$${}_n\mathrm{H}_{r-n}$$

이다.

(2) 조건을 만족시키는 함수의 개수

공집합이 아닌 두 집합 $X$, $Y$의 원소의 개수가 각각 $m$, $n$일 때, 집합 $X$에서 집합 $Y$로의 함수 중에서

'집합 $X$의 임의의 두 원소 $x_1$, $x_2$에 대하여 $x_1<x_2$이면 $f(x_1)\le f(x_2)$이다.'

를 만족시키는 함수 $f$의 개수는

$${}_n\mathrm{H}_m$$

이다.

**설명** 위의 조건을 만족시키는 함수는 집합 $Y$의 원소 $n$개에서 중복을 허락하여 $m$개를 택하여 집합 $X$의 원소에 크기순으로 대응시키면 되므로 구하는 함수의 개수는 서로 다른 $n$개에서 $m$개를 택하는 중복조합의 수 ${}_n\mathrm{H}_m$과 같다.

## 3. 이항정리

(1) 이항정리

자연수 $n$에 대하여 다항식 $(a+b)^n$을 전개하면

$$(a+b)^n={}_n\mathrm{C}_0a^n+{}_n\mathrm{C}_1a^{n-1}b+{}_n\mathrm{C}_2a^{n-2}b^2+\cdots+{}_n\mathrm{C}_ra^{n-r}b^r+\cdots+{}_n\mathrm{C}_nb^n$$

이다. 이와 같이 다항식 $(a+b)^n$을 전개하는 것을 이항정리라고 한다.

**설명** 다항식 $(a+b)^3$을 전개하면

$$(a+b)^3=(a+b)(a+b)(a+b)=a^3+3a^2b+3ab^2+b^3$$

이때 $a^2b$항은 세 개의 인수 $(a+b)$ 중 어느 한 인수에서 $b$를 택하고, 나머지 두 인수에서 각각 $a$를 택하여 곱한 단항식 $baa$, $aba$, $aab$의 합이다.

즉, $a^2b$의 계수는 세 개의 인수 $(a+b)$ 중 한 개에서 $b$를 택하는 조합의 수와 같으므로 ${}_3\mathrm{C}_1=3$이다.

| $(a+b)$ | $(a+b)$ | $(a+b)$ | | |
|---|---|---|---|---|
| ⬇ | ⬇ | ⬇ | | |
| $b$ | $a$ | $a$ | ➡ | $a^2b$ |
| $a$ | $b$ | $a$ | ➡ | $a^2b$ |
| $a$ | $a$ | $b$ | ➡ | $a^2b$ |

마찬가지 방법으로 $a^3$, $ab^2$, $b^3$의 계수는 각각 ${}_3\mathrm{C}_0$, ${}_3\mathrm{C}_2$, ${}_3\mathrm{C}_3$임을 알 수 있다.

따라서 $(a+b)^3$의 전개식을 조합의 수를 이용하여 나타내면

$$(a+b)^3={}_3\mathrm{C}_0a^3+{}_3\mathrm{C}_1a^2b+{}_3\mathrm{C}_2ab^2+{}_3\mathrm{C}_3b^3$$

이다.

일반적으로 자연수 $n$에 대하여 다항식

$$(a+b)^n=\underbrace{(a+b)(a+b)\times\cdots\times(a+b)}_{n\text{개}}$$

의 전개식에서 $a^{n-r}b^r$항은 $n$개의 인수 $(a+b)$ 중 $r$개의 인수에서 $b$를 택하고, 나머지 $(n-r)$개의 인수에서 $a$를 택하여 곱한 것이므로 $a^{n-r}b^r$의 계수는 $n$개의 인수 $(a+b)$ 중 $r$개의 인수에서 $b$를 택하는 조합의 수와 같다.

즉, 다항식 $(a+b)^n$의 전개식에서 $a^{n-r}b^r$의 계수는 ${}_n\mathrm{C}_r$와 같다.

따라서 다항식 $(a+b)^n$의 전개식은

$$(a+b)^n={}_n\mathrm{C}_0a^n+{}_n\mathrm{C}_1a^{n-1}b+{}_n\mathrm{C}_2a^{n-2}b^2+\cdots+{}_n\mathrm{C}_ra^{n-r}b^r+\cdots+{}_n\mathrm{C}_nb^n$$

**참고** ${}_n\mathrm{C}_r={}_n\mathrm{C}_{n-r}$이므로 다항식 $(a+b)^n$의 전개식에서 $a^{n-r}b^r$의 계수와 $a^rb^{n-r}$의 계수는 같다.

(2) 이항계수

다항식 $(a+b)^n$의 전개식에서 각 항의 계수

$${}_n\mathrm{C}_0,\ {}_n\mathrm{C}_1,\ {}_n\mathrm{C}_2,\ \cdots,\ {}_n\mathrm{C}_r,\ \cdots,\ {}_n\mathrm{C}_n$$

을 이항계수라 하고, ${}_n\mathrm{C}_ra^{n-r}b^r$ $(r=0,\ 1,\ 2,\ \cdots,\ n)$을 일반항이라고 한다.

**예** 다항식 $(x+2y)^5$의 전개식의 일반항을 이용하여 $x^3y^2$의 계수를 구해 보자.

다항식 $(x+2y)^5$의 전개식의 일반항은

$${}_5\mathrm{C}_rx^{5-r}(2y)^r={}_5\mathrm{C}_r2^rx^{5-r}y^r\ (r=0,\ 1,\ 2,\ 3,\ 4,\ 5)$$

이므로 $x^3y^2$항은 $r=2$일 때이다.

따라서 $x^3y^2$의 계수는

$${}_5\mathrm{C}_2\times2^2=10\times4=40$$

# 03 확률의 뜻과 활용

## 1. 시행과 사건

(1) 시행

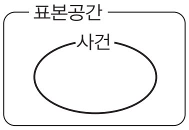
주사위나 동전 던지기와 같이 같은 조건에서 반복할 수 있고, 그 결과가 우연에 의하여 정해지는 실험이나 관찰을 시행이라고 한다.

(2) 사건

① 표본공간: 어떤 시행에서 일어날 수 있는 모든 결과의 집합을 표본공간이라고 한다.

② 사건: 표본공간의 부분집합을 사건이라고 한다.

③ 근원사건: 한 개의 원소로 이루어진 사건을 근원사건이라고 한다.

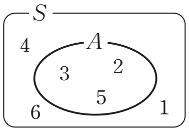



예 한 개의 주사위를 한 번 던져서 나오는 눈의 수를 확인하는 시행에서

① 표본공간을 $S$라 하면 $S=\{1, 2, 3, 4, 5, 6\}$

② 소수의 눈이 나오는 사건을 $A$라 하면 $A=\{2, 3, 5\}$

③ 근원사건은 $\{1\}$, $\{2\}$, $\{3\}$, $\{4\}$, $\{5\}$, $\{6\}$




참고 표본공간은 공집합이 아닌 경우만 생각한다.

## 2. 배반사건과 여사건

표본공간 $S$의 두 사건 $A$, $B$에 대하여

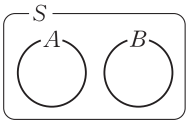
(1) 사건 $A$ 또는 사건 $B$가 일어나는 사건을 $A\cup B$와 같이 나타낸다.

(2) 두 사건 $A$와 $B$가 동시에 일어나는 사건을 $A\cap B$와 같이 나타낸다.

(3) 배반사건: 두 사건 $A$와 $B$가 동시에 일어나지 않을 때, 즉

$$A\cap B=\varnothing$$

일 때, 두 사건 $A$와 $B$는 서로 배반사건이라고 한다.

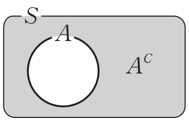



(4) 여사건: 사건 $A$에 대하여 사건 $A$가 일어나지 않는 사건을 $A$의 여사건이라 하고, 기호로

$$A^C$$

과 같이 나타낸다.

이때 $A\cap A^C=\varnothing$이므로 사건 $A$와 그 여사건 $A^C$은 서로 배반사건이다.




예 한 개의 주사위를 한 번 던지는 시행에서 6의 약수의 눈이 나오는 사건을 $A$, 홀수의 눈이 나오는 사건을 $B$, 4의 배수의 눈이 나오는 사건을 $C$라 하면

$$A=\{1, 2, 3, 6\},\ B=\{1, 3, 5\},\ C=\{4\}$$

① 사건 $A$ 또는 사건 $B$가 일어나는 사건은 $A\cup B=\{1, 2, 3, 5, 6\}$

② 두 사건 $A$와 $B$가 동시에 일어나는 사건은 $A\cap B=\{1, 3\}$

③ $A\cap C=\varnothing$이므로 두 사건 $A$와 $C$는 서로 배반사건이다.

④ 사건 $A$의 여사건은 $A^C=\{4, 5\}$이다.

## 3. 확률의 뜻

(1) 확률

어떤 시행에서 사건 $A$가 일어날 가능성을 수로 나타낸 것을 사건 $A$의 확률이라 하고, 기호로

$$\mathrm{P}(A)$$

와 같이 나타낸다.

(2) 수학적 확률

표본공간이 $S$인 어떤 시행에서 각각의 근원사건이 일어날 가능성이 모두 같은 정도로 기대될 때, 표본공간 $S$의 사건 $A$가 일어날 확률 $\mathrm{P}(A)$를

$$\mathrm{P}(A)=\frac{n(A)}{n(S)}=\frac{(\text{사건 } A\text{의 원소의 개수})}{(\text{표본공간 } S\text{의 원소의 개수})}$$

로 정의하고, 이것을 사건 $A$가 일어날 수학적 확률이라고 한다.

**예** 서로 다른 2개의 주사위를 동시에 한 번 던질 때, 나오는 두 눈의 수의 합이 6이 될 확률과 곱이 12가 될 확률을 각각 구해 보자.

서로 다른 2개의 주사위를 동시에 한 번 던지는 시행에서 표본공간을 $S$라 하면

$S=\{(1, 1), (1, 2), \cdots, (6, 5), (6, 6)\}$이므로 $n(S)=6\times6=36$

이때 나오는 두 눈의 수의 합이 6인 사건을 $A$라 하면

$A=\{(1, 5), (2, 4), (3, 3), (4, 2), (5, 1)\}$이므로 $n(A)=5$

따라서 $\mathrm{P}(A)=\dfrac{n(A)}{n(S)}=\dfrac{5}{36}$

한편, 나오는 두 눈의 수의 곱이 12인 사건을 $B$라 하면

$B=\{(2, 6), (3, 4), (4, 3), (6, 2)\}$이므로 $n(B)=4$

따라서 $\mathrm{P}(B)=\dfrac{n(B)}{n(S)}=\dfrac{4}{36}=\dfrac{1}{9}$

**참고** 수학적 확률은 표본공간이 공집합이 아닌 유한집합인 경우에만 생각한다.

(3) 통계적 확률

같은 시행을 $n$번 반복할 때, 사건 $A$가 일어난 횟수를 $r_n$이라 하자. 이때 시행 횟수 $n$이 한없이 커짐에 따라 상대도수 $\dfrac{r_n}{n}$이 일정한 값 $p$에 가까워질 때, 이 값 $p$를 사건 $A$가 일어날 통계적 확률이라고 한다. 그러나 실제로 $n$의 값을 한없이 크게 할 수 없으므로 $n$이 충분히 클 때의 상대도수 $\dfrac{r_n}{n}$을 통계적 확률로 생각한다.

**참고** 일반적으로 사건 $A$가 일어날 수학적 확률이 $p$일 때, 시행 횟수 $n$을 충분히 크게 하면 사건 $A$가 일어나는 상대도수 $\dfrac{r_n}{n}$은 수학적 확률 $p$에 가까워진다.

# 03 확률의 뜻과 활용

## 4. 확률의 기본 성질

표본공간 $S$가 유한개의 근원사건으로 이루어져 있는 어떤 시행에서

(1) 임의의 사건 $A$에 대하여 $0 \le \mathrm{P}(A) \le 1$

(2) 반드시 일어나는 사건 $S$에 대하여 $\mathrm{P}(S)=1$

(3) 절대로 일어나지 않는 사건 $\varnothing$에 대하여 $\mathrm{P}(\varnothing)=0$

참고 어떤 시행에서 표본공간 $S$의 임의의 사건 $A$에 대하여

$\varnothing \subset A \subset S$이므로 $0 \le n(A) \le n(S)$

이 부등식의 각 변을 $n(S)$로 나누면

$0 \le \dfrac{n(A)}{n(S)} \le 1$, 즉 $0 \le \mathrm{P}(A) \le 1$

특히 반드시 일어나는 사건 $S$에 대하여 $\mathrm{P}(S)=\dfrac{n(S)}{n(S)}=1$이고,

절대로 일어나지 않는 사건 $\varnothing$에 대하여 $n(\varnothing)=0$이므로 $\mathrm{P}(\varnothing)=\dfrac{n(\varnothing)}{n(S)}=\dfrac{0}{n(S)}=0$이다.

예 흰 공 3개와 검은 공 2개가 들어 있는 주머니에서 임의로 3개의 공을 동시에 꺼낼 때, 흰 공이 1개 이상 나올 확률은 1이고, 검은 공이 3개 나올 확률은 0이다.

## 5. 확률의 덧셈정리

표본공간 $S$의 두 사건 $A$, $B$에 대하여 사건 $A$ 또는 사건 $B$가 일어날 확률은

$$\mathrm{P}(A\cup B)=\mathrm{P}(A)+\mathrm{P}(B)-\mathrm{P}(A\cap B)$$

특히 두 사건 $A$와 $B$가 서로 배반사건이면

$$\mathrm{P}(A\cup B)=\mathrm{P}(A)+\mathrm{P}(B)$$

설명 표본공간 $S$의 두 사건 $A$, $B$에 대하여

$$n(A\cup B)=n(A)+n(B)-n(A\cap B)$$

이 식의 양변을 $n(S)$로 나누면

$$\frac{n(A\cup B)}{n(S)}=\frac{n(A)}{n(S)}+\frac{n(B)}{n(S)}-\frac{n(A\cap B)}{n(S)}$$

$$\mathrm{P}(A\cup B)=\mathrm{P}(A)+\mathrm{P}(B)-\mathrm{P}(A\cap B)$$

특히 두 사건 $A$와 $B$가 서로 배반사건, 즉 $A\cap B=\varnothing$이면 $\mathrm{P}(A\cap B)=0$이므로

$$\mathrm{P}(A\cup B)=\mathrm{P}(A)+\mathrm{P}(B)$$

예 1부터 10까지의 자연수가 하나씩 적힌 10장의 카드 중에서 임의로 한 장의 카드를 선택할 때, 선택한 카드에 적힌 수가 2의 배수이거나 3의 배수일 확률을 구해 보자.

임의로 카드 한 장을 선택할 때, 선택한 카드에 적힌 수가 2의 배수인 사건을 $A$, 3의 배수인 사건을 $B$라 하면 사건 $A\cap B$는 6의 배수인 사건이므로

$$\mathrm{P}(A)=\frac{1}{2},\ \mathrm{P}(B)=\frac{3}{10},\ \mathrm{P}(A\cap B)=\frac{1}{10}$$

따라서 구하는 확률은 확률의 덧셈정리에 의하여

$$\mathrm{P}(A\cup B)=\mathrm{P}(A)+\mathrm{P}(B)-\mathrm{P}(A\cap B)=\frac{1}{2}+\frac{3}{10}-\frac{1}{10}=\frac{7}{10}$$

# 04 조건부확률

## 1. 조건부확률

(1) 조건부확률의 뜻

표본공간 $S$의 두 사건 $A$, $B$에 대하여 확률이 0이 아닌 사건 $A$가 일어났을 때, 사건 $B$가 일어날 확률을 사건 $A$가 일어났을 때의 사건 $B$의 조건부확률이라 하고, 기호로

$$\mathrm{P}(B|A)$$

와 같이 나타낸다.

(2) 조건부확률의 계산

사건 $A$가 일어났을 때의 사건 $B$의 조건부확률은

$$\mathrm{P}(B|A)=\frac{\mathrm{P}(A\cap B)}{\mathrm{P}(A)} \text{ (단, } \mathrm{P}(A)>0\text{)}$$

**설명** 표본공간이 $S$인 어떤 시행에서 각 근원사건이 일어날 가능성이 같은 정도로 기대될 때, 표본공간 $S$의 두 사건 $A$, $B$에 대하여 조건부확률 $\mathrm{P}(B|A)$는 사건 $A$를 새로운 표본공간으로 하여 사건 $B$, 즉 사건 $A\cap B$가 일어날 확률이므로

$$\mathrm{P}(B|A)=\frac{n(A\cap B)}{n(A)}$$

이 등식의 우변의 분모와 분자를 각각 $n(S)$로 나누면

$$\mathrm{P}(B|A)=\frac{n(A\cap B)}{n(A)}=\frac{\frac{n(A\cap B)}{n(S)}}{\frac{n(A)}{n(S)}}=\frac{\mathrm{P}(A\cap B)}{\mathrm{P}(A)}$$

**예** 한 개의 주사위를 한 번 던져서 나온 눈의 수가 소수일 때, 그 수가 홀수일 확률을 구해 보자.

한 개의 주사위를 한 번 던지는 시행에서 표본공간을 $S$, 소수의 눈이 나오는 사건을 $A$, 홀수의 눈이 나오는 사건을 $B$라 하면

$S=\{1, 2, 3, 4, 5, 6\}$, $A=\{2, 3, 5\}$, $B=\{1, 3, 5\}$, $A\cap B=\{3, 5\}$이므로

$$\mathrm{P}(A)=\frac{n(A)}{n(S)}=\frac{3}{6}=\frac{1}{2}$$

$$\mathrm{P}(A\cap B)=\frac{n(A\cap B)}{n(S)}=\frac{2}{6}=\frac{1}{3}$$

따라서 사건 $A$가 일어났을 때의 사건 $B$의 조건부확률은

$$\mathrm{P}(B|A)=\frac{\mathrm{P}(A\cap B)}{\mathrm{P}(A)}=\frac{\frac{1}{3}}{\frac{1}{2}}=\frac{2}{3}$$

[다른 풀이] $n(A)=3$, $n(A\cap B)=2$이므로

$$\mathrm{P}(B|A)=\frac{n(A\cap B)}{n(A)}=\frac{2}{3}$$

## 2. 확률의 곱셈정리

(1) 확률의 곱셈정리

두 사건 $A$, $B$에 대하여

$$\mathrm{P}(A\cap B)=\mathrm{P}(A)\mathrm{P}(B|A)=\mathrm{P}(B)\mathrm{P}(A|B) \text{ (단, } \mathrm{P}(A)>0,\ \mathrm{P}(B)>0\text{)}$$

설명 $\mathrm{P}(A)>0$일 때, 사건 $A$가 일어났을 때의 사건 $B$의 조건부확률은

$$\mathrm{P}(B|A)=\frac{\mathrm{P}(A\cap B)}{\mathrm{P}(A)} \quad \cdots\cdots ㉠$$

이므로 ㉠의 양변에 $\mathrm{P}(A)$를 곱하면

$$\mathrm{P}(A\cap B)=\mathrm{P}(A)\mathrm{P}(B|A)$$

마찬가지로 $\mathrm{P}(B)>0$일 때, 사건 $B$가 일어났을 때의 사건 $A$의 조건부확률은

$$\mathrm{P}(A|B)=\frac{\mathrm{P}(A\cap B)}{\mathrm{P}(B)} \quad \cdots\cdots ㉡$$

이므로 ㉡의 양변에 $\mathrm{P}(B)$를 곱하면

$$\mathrm{P}(A\cap B)=\mathrm{P}(B)\mathrm{P}(A|B)$$

따라서 $\mathrm{P}(A\cap B)=\mathrm{P}(A)\mathrm{P}(B|A)=\mathrm{P}(B)\mathrm{P}(A|B)$가 성립한다.

예 흰 공 6개와 검은 공 3개가 들어 있는 주머니에서 임의로 공을 한 개씩 두 번 꺼낸다. 꺼낸 공을 주머니에 다시 넣지 않을 때, 꺼낸 공이 모두 흰 공일 확률을 구해 보자.

첫 번째 꺼낸 공이 흰 공인 사건을 $A$, 두 번째 꺼낸 공이 흰 공인 사건을 $B$라 하면

$$\mathrm{P}(A)=\frac{6}{9}=\frac{2}{3},\ \mathrm{P}(B|A)=\frac{5}{8}$$

따라서 구하는 확률은

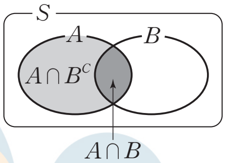
$$\mathrm{P}(A\cap B)=\mathrm{P}(A)\mathrm{P}(B|A)=\frac{2}{3}\times\frac{5}{8}=\frac{5}{12}$$

(2) 확률의 곱셈정리의 활용

두 사건 $A$, $B$에 대하여

$$\mathrm{P}(A)=\mathrm{P}(B)\mathrm{P}(A|B)+\mathrm{P}(B^C)\mathrm{P}(A|B^C) \text{ (단, } 0<\mathrm{P}(B)<1\text{)}$$

설명 두 사건 $A$, $B$에 대하여




$$A=(A\cap B)\cup(A\cap B^C)$$

이고, 두 사건 $A\cap B$와 $A\cap B^C$은 서로 배반사건이므로 확률의 덧셈정리에 의하여

$$\mathrm{P}(A)=\mathrm{P}(A\cap B)+\mathrm{P}(A\cap B^C)$$

이때 $0<\mathrm{P}(B)<1$이면 $0<\mathrm{P}(B^C)<1$이므로 확률의 곱셈정리에 의하여

$$\mathrm{P}(A\cap B)=\mathrm{P}(B)\mathrm{P}(A|B),\ \mathrm{P}(A\cap B^C)=\mathrm{P}(B^C)\mathrm{P}(A|B^C)$$

따라서 $\mathrm{P}(A)=\mathrm{P}(B)\mathrm{P}(A|B)+\mathrm{P}(B^C)\mathrm{P}(A|B^C)$이 성립한다.

# 04 조건부확률

## 3. 사건의 독립과 종속

(1) 사건의 독립

두 사건 $A$, $B$에 대하여 $\mathrm{P}(A)>0$, $\mathrm{P}(B)>0$이고, 어느 한 사건이 일어나는 것이 다른 사건이 일어날 확률에 영향을 주지 않을 때, 즉

$\mathrm{P}(B|A)=\mathrm{P}(B)$ 또는 $\mathrm{P}(A|B)=\mathrm{P}(A)$

일 때, 두 사건 $A$와 $B$는 서로 독립이라고 한다.

(2) 사건의 종속

두 사건 $A$와 $B$가 서로 독립이 아닐 때, 두 사건 $A$와 $B$는 서로 종속이라고 한다.

(3) 두 사건이 서로 독립일 조건

두 사건 $A$와 $B$가 서로 독립이기 위한 필요충분조건은

$\mathrm{P}(A\cap B)=\mathrm{P}(A)\mathrm{P}(B)$ (단, $\mathrm{P}(A)>0$, $\mathrm{P}(B)>0$)

**설명** 두 사건 $A$와 $B$가 서로 독립이면 $\mathrm{P}(B|A)=\mathrm{P}(B)$이므로 확률의 곱셈정리에 의하여

$$\mathrm{P}(A\cap B)=\mathrm{P}(A)\mathrm{P}(B|A)=\mathrm{P}(A)\mathrm{P}(B)$$

가 성립한다.

역으로 $\mathrm{P}(A\cap B)=\mathrm{P}(A)\mathrm{P}(B)$이면

$$\mathrm{P}(B|A)=\frac{\mathrm{P}(A\cap B)}{\mathrm{P}(A)}=\frac{\mathrm{P}(A)\mathrm{P}(B)}{\mathrm{P}(A)}=\mathrm{P}(B)$$

이므로 두 사건 $A$와 $B$는 서로 독립이다.

**참고** $0<\mathrm{P}(A)<1$, $0<\mathrm{P}(B)<1$인 두 사건 $A$와 $B$가 서로 독립이면

$$\mathrm{P}(A\cap B^C)=\mathrm{P}(A)-\mathrm{P}(A\cap B)=\mathrm{P}(A)-\mathrm{P}(A)\mathrm{P}(B)=\mathrm{P}(A)\{1-\mathrm{P}(B)\}=\mathrm{P}(A)\mathrm{P}(B^C)$$

이므로 두 사건 $A$와 $B^C$은 서로 독립이다.

마찬가지로 두 사건 $A^C$과 $B$는 서로 독립이고, 두 사건 $A^C$과 $B^C$도 서로 독립이다.

**예** 한 개의 주사위를 한 번 던질 때, 3의 배수의 눈이 나오는 사건을 $A$, 짝수의 눈이 나오는 사건을 $B$, 5 이상의 눈이 나오는 사건을 $C$라 할 때, 두 사건 $A$와 $B$, 두 사건 $A$와 $C$가 서로 독립인지 종속인지를 각각 알아보자.

$A=\{3, 6\}$, $B=\{2, 4, 6\}$, $C=\{5, 6\}$이므로

$$\mathrm{P}(A)=\frac{2}{6}=\frac{1}{3},\ \mathrm{P}(B)=\frac{3}{6}=\frac{1}{2},\ \mathrm{P}(C)=\frac{2}{6}=\frac{1}{3}$$

① $\mathrm{P}(A)\mathrm{P}(B)=\frac{1}{3}\times\frac{1}{2}=\frac{1}{6}$이고, $A\cap B=\{6\}$에서 $\mathrm{P}(A\cap B)=\frac{1}{6}$

따라서 $\mathrm{P}(A\cap B)=\mathrm{P}(A)\mathrm{P}(B)$이므로 두 사건 $A$와 $B$는 서로 독립이다.

② $\mathrm{P}(A)\mathrm{P}(C)=\frac{1}{3}\times\frac{1}{3}=\frac{1}{9}$이고, $A\cap C=\{6\}$에서 $\mathrm{P}(A\cap C)=\frac{1}{6}$

따라서 $\mathrm{P}(A\cap C)\neq\mathrm{P}(A)\mathrm{P}(C)$이므로 두 사건 $A$와 $C$는 서로 종속이다.

# 05 이산확률변수의 확률분포

## 1. 확률변수

(1) 확률변수: 어떤 시행에서 표본공간의 각 원소에 하나의 실수의 값을 대응시키는 함수를 확률변수라고 한다. 확률변수 $X$가 어떤 값 $x$를 가질 확률을 기호로 $\mathrm{P}(X=x)$와 같이 나타낸다.

예 두 개의 동전을 동시에 던지는 시행에서 앞면이 나오는 동전의 개수를 확률변수 $X$라 하면 $X$가 갖는 값은 0, 1, 2이다.

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

(2) 이산확률변수: 확률변수 $X$가 갖는 값이 유한개이거나 무한히 많더라도 자연수와 같이 셀 수 있을 때, 그 확률변수 $X$를 이산확률변수라고 한다.

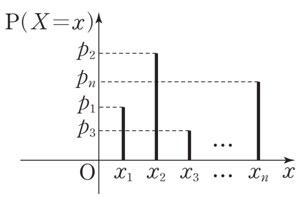
예 한 개의 주사위를 두 번 던지는 시행에서 나온 두 눈의 수의 합을 확률변수 $X$라 할 때, $X$가 갖는 값은 2부터 12까지의 자연수로 유한개이므로 $X$는 이산확률변수이다.

## 2. 이산확률변수의 확률분포

(1) 이산확률변수의 확률분포: 이산확률변수 $X$가 갖는 값이 $x_1$, $x_2$, $x_3$, $\cdots$, $x_n$이고 $X$가 이 값들을 가질 확률이 각각 $p_1$, $p_2$, $p_3$, $\cdots$, $p_n$일 때, $x_1$, $x_2$, $x_3$, $\cdots$, $x_n$과 $p_1$, $p_2$, $p_3$, $\cdots$, $p_n$ 사이의 대응 관계를 이산확률변수 $X$의 확률분포라고 한다.

이때 이산확률변수 $X$의 확률분포는 다음과 같이 표 또는 그래프로 나타낼 수 있다.

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




(2) 확률질량함수: 이산확률변수 $X$가 갖는 값 $x_i$ $(i=1, 2, 3, \cdots, n)$과 $X$가 이 값들을 가질 확률 $p_i$ $(i=1, 2, 3, \cdots, n)$ 사이의 대응 관계를 나타내는 함수

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

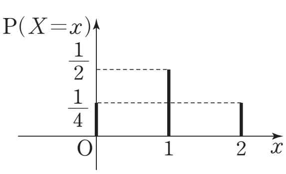
을 이산확률변수 $X$의 확률질량함수라고 한다.

예 한 개의 동전을 두 번 던지는 시행에서 앞면이 나오는 횟수를 확률변수 $X$라 하면 $X$가 갖는 값은 0, 1, 2이므로 $X$는 이산확률변수이고, $X$의 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 $X$의 확률분포를 표와 그래프로 나타내면 다음과 같다.

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



## 3. 확률질량함수의 성질

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

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

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

예 한 개의 주사위를 두 번 던지는 시행에서 3의 배수의 눈이 나오는 횟수를 확률변수 $X$라 하면 $X$가 갖는 값이 0, 1, 2이고 $X$의 확률질량함수는

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

이므로

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

따라서

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

이므로 확률질량함수가 위의 성질 (1), (2)를 만족시킴을 확인할 수 있다.

## 4. 이산확률변수 $X$의 기댓값(평균)

이산확률변수 $X$의 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 |

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

를 확률변수 $X$의 기댓값 또는 평균이라 하고, 기호로

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

와 같이 나타낸다.

예 한 개의 동전을 세 번 던지는 시행에서 앞면이 나오는 횟수를 확률변수 $X$라 할 때, $X$의 기댓값 $\mathrm{E}(X)$를 구해 보자.
확률변수 $X$가 갖는 값이 0, 1, 2, 3이고 $X$의 확률질량함수는

$$\mathrm{P}(X=x)={}_3\mathrm{C}_x\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{3-x}=\frac{{}_3\mathrm{C}_x}{8}\ (x=0, 1, 2, 3)$$

이다. 이때 확률변수 $X$의 확률분포를 표로 나타내면 다음과 같다.

| $X$ | 0 | 1 | 2 | 3 | 합계 |
|---|---|---|---|---|---|
| $\mathrm{P}(X=x)$ | $\frac{1}{8}$ | $\frac{3}{8}$ | $\frac{3}{8}$ | $\frac{1}{8}$ | 1 |

따라서 $\mathrm{E}(X)=0\times\frac{1}{8}+1\times\frac{3}{8}+2\times\frac{3}{8}+3\times\frac{1}{8}=\frac{12}{8}=\frac{3}{2}$

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

## 5. 이산확률변수 $X$의 분산, 표준편차

이산확률변수 $X$의 확률분포가 오른쪽 표와 같을 때, 확률변수 $X$의 분산과 표준편차는 다음과 같다.

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

(1) 분산

$\mathrm{E}(X)=m$일 때, $(X-m)^2$의 평균

$$\mathrm{E}((X-m)^2)=\sum_{i=1}^{n}(x_i-m)^2p_i$$
$$=(x_1-m)^2p_1+(x_2-m)^2p_2+(x_3-m)^2p_3+\cdots+(x_n-m)^2p_n$$

을 확률변수 $X$의 분산이라 하고, 기호로

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

와 같이 나타낸다. 이때

$$\mathrm{V}(X)=\sum_{i=1}^{n}(x_i-m)^2p_i$$
$$=\sum_{i=1}^{n}(x_i^2p_i-2mx_ip_i+m^2p_i)=\sum_{i=1}^{n}x_i^2p_i-2m\sum_{i=1}^{n}x_ip_i+m^2\sum_{i=1}^{n}p_i$$
$$=\sum_{i=1}^{n}x_i^2p_i-2m^2+m^2\ \left(\sum_{i=1}^{n}x_ip_i=m,\ \sum_{i=1}^{n}p_i=1\text{이므로}\right)$$
$$=\sum_{i=1}^{n}x_i^2p_i-m^2$$

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

(2) 표준편차

분산 $\mathrm{V}(X)$의 양의 제곱근을 확률변수 $X$의 표준편차라 하고, 기호로 $\sigma(X)$와 같이 나타낸다.

즉, $\sigma(X)=\sqrt{\mathrm{V}(X)}$

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

**예** 확률변수 $X$의 확률분포가 오른쪽 표와 같을 때, $X$의 분산, 표준편차를 구해 보자.

| $X$ | 1 | 2 | 3 | 합계 |
|---|---|---|---|---|
| $\mathrm{P}(X=x)$ | $\frac{1}{2}$ | $\frac{1}{3}$ | $\frac{1}{6}$ | 1 |

$$m=\mathrm{E}(X)=1\times\frac{1}{2}+2\times\frac{1}{3}+3\times\frac{1}{6}=\frac{5}{3}$$

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

$$\mathrm{V}(X)=\left(1-\frac{5}{3}\right)^2\times\frac{1}{2}+\left(2-\frac{5}{3}\right)^2\times\frac{1}{3}+\left(3-\frac{5}{3}\right)^2\times\frac{1}{6}=\frac{5}{9}$$
$$\sigma(X)=\sqrt{\mathrm{V}(X)}=\sqrt{\frac{5}{9}}=\frac{\sqrt{5}}{3}$$

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

$$\mathrm{E}(X^2)=1^2\times\frac{1}{2}+2^2\times\frac{1}{3}+3^2\times\frac{1}{6}=\frac{10}{3}$$
$$\mathrm{V}(X)=\mathrm{E}(X^2)-\{\mathrm{E}(X)\}^2=\frac{10}{3}-\left(\frac{5}{3}\right)^2=\frac{5}{9}$$
$$\sigma(X)=\sqrt{\mathrm{V}(X)}=\sqrt{\frac{5}{9}}=\frac{\sqrt{5}}{3}$$

### 6. 이산확률변수 $aX+b$의 평균, 분산, 표준편차

이산확률변수 $X$와 두 상수 $a$, $b$ $(a\neq 0)$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) $\sigma(aX+b)=|a|\sigma(X)$

설명 이산확률변수 $X$의 확률분포가 다음과 같을 때, 이산확률변수

$Y=aX+b$ ($a$, $b$는 상수, $a\neq 0$)

의 평균, 분산, 표준편차를 구해 보자.

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

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

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

이므로 확률변수 $Y$의 확률분포를 표로 나타내면 다음과 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 $Y$의 평균은

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

확률변수 $X$의 평균을 $m$이라 하면 확률변수 $Y$의 평균은 $am+b$이므로 $Y$의 분산과 표준편차는

$$\begin{aligned}\mathrm{V}(Y)&=\sum_{i=1}^{n}\{y_i-(am+b)\}^2p_i\\&=\sum_{i=1}^{n}\{(ax_i+b)-(am+b)\}^2p_i\\&=\sum_{i=1}^{n}a^2(x_i-m)^2p_i\\&=a^2\sum_{i=1}^{n}(x_i-m)^2p_i=a^2\mathrm{V}(X)\end{aligned}$$

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

예 이산확률변수 $X$에 대하여 $\mathrm{E}(X)=6$, $\mathrm{V}(X)=9$일 때, 확률변수 $2X+1$의 평균, 분산, 표준편차는 다음과 같다.

$\mathrm{E}(2X+1)=2\mathrm{E}(X)+1=2\times 6+1=13$

$\mathrm{V}(2X+1)=2^2\mathrm{V}(X)=4\times 9=36$

$\sigma(2X+1)=|2|\sigma(X)=2\times\sqrt{9}=6$

## 7. 이항분포

한 번의 시행에서 사건 $A$가 일어날 확률이 $p$로 일정할 때, $n$번의 독립시행에서 사건 $A$가 일어나는 횟수를 확률변수 $X$라 하면 $X$가 갖는 값은 0, 1, 2, $\cdots$, $n$이고, $X$의 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 $X$의 확률분포를 이항분포라 하고, 기호로 $\mathrm{B}(n,\ p)$와 같이 나타낸다.
이때 확률변수 $X$는 이항분포 $\mathrm{B}(n,\ p)$를 따른다고 하며, $X$의 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 |

**참고** (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 $\sum\limits_{x=0}^{n} {}_n\mathrm{C}_x p^x q^{n-x}=1$임을 알 수 있다.

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

**예** 한 개의 주사위를 한 번 던지는 시행에서 4의 약수의 눈이 나오는 사건을 $A$라 하고, 이 주사위를 10번 던질 때 사건 $A$가 일어나는 횟수를 확률변수 $X$라 하자. 주사위를 한 번 던질 때 사건 $A$가 일어날 확률이 $\frac{1}{2}$이고 독립시행의 횟수가 10이므로 확률변수 $X$는 이항분포 $\mathrm{B}\left(10,\ \frac{1}{2}\right)$을 따른다.

## 8. 이항분포의 평균, 분산, 표준편차

확률변수 $X$가 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, $q=1-p$)

## 9. 큰수의 법칙

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

**참고** 큰수의 법칙에 의하여 시행 횟수 $n$이 충분히 클 때, 사건 $A$의 상대도수는 수학적 확률에 가까워지므로 사건 $A$의 상대도수 $\frac{X}{n}$를 사건 $A$가 일어날 확률 $\mathrm{P}(A)$로 간주할 수 있다. 따라서 자연현상이나 사회현상에서 수학적 확률을 구하기 어려운 경우에는 시행 횟수를 충분히 크게 한 후 사건의 상대도수를 구하여 수학적 확률로 이용할 수 있다.

# 06 연속확률변수의 확률분포

## 1. 연속확률변수

확률변수 $X$가 어떤 구간에 속하는 모든 실수의 값을 가질 때, $X$를 연속확률변수라고 한다.

**참고** 길이, 무게, 온도, 시간 등의 값을 확률변수 $X$라 하면 $X$는 어떤 범위에 속하는 모든 실수의 값을 갖는다.

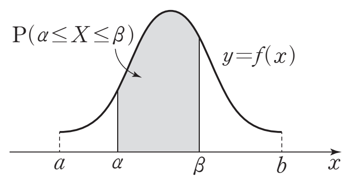
## 2. 확률밀도함수

일반적으로 $a \le X \le b$의 모든 실수의 값을 가지는 연속확률변수 $X$에 대하여 $a \le x \le b$에서 정의된 함수 $f(x)$가 다음의 세 가지를 모두 만족시킬 때, 함수 $f(x)$를 연속확률변수 $X$의 확률밀도함수라고 한다.

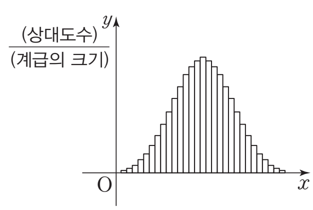
① $f(x) \ge 0$ (단, $a \le x \le b$)

② 함수 $y=f(x)$의 그래프와 $x$축 및 두 직선 $x=a$, $x=b$로 둘러싸인 부분의 넓이는 1이다.

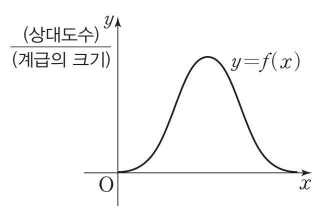
③ $\mathrm{P}(\alpha \le X \le \beta)$는 함수 $y=f(x)$의 그래프와 $x$축 및 두 직선 $x=\alpha$, $x=\beta$로 둘러싸인 부분의 넓이와 같다. (단, $a \le \alpha \le \beta \le b$)




**설명** 연속확률변수 $X$의 $\dfrac{(\text{상대도수})}{(\text{계급의 크기})}$를 히스토그램으로 나타내면 히스토그램의 각 구간에 세워진 직사각형의 넓이는 각 구간의 상대도수를 나타내고, 상대도수의 합이 1이므로 모든 직사각형들의 넓이의 합은 항상 1이다.
이때 조사 대상의 수를 한없이 늘리고, 계급의 크기를 0에 가깝게 하여 히스토그램을 그리면 [그림 1]과 같이 어떤 곡선 모양에 가까워지고, 이 과정을 계속하면 [그림 2]와 같이 매끄러운 곡선이 된다.




[그림 1]




[그림 2]

**참고** 연속확률변수 $X$가 하나의 값을 가질 확률은 0이다.
즉, $\mathrm{P}(X=\alpha)=\mathrm{P}(X=\beta)=0$이므로

$$\begin{aligned}\mathrm{P}(\alpha < X < \beta) &= \mathrm{P}(\alpha \le X < \beta)\\ &= \mathrm{P}(\alpha < X \le \beta)\\ &= \mathrm{P}(\alpha \le X \le \beta)\end{aligned}$$

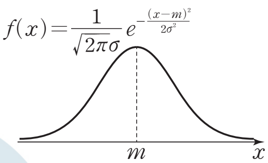
## 3. 정규분포

연속확률변수 $X$가 모든 실수의 값을 갖고, 그 확률밀도함수 $f(x)$가 두 상수 $m$, $\sigma$ $(\sigma>0)$에 대하여

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-m)^2}{2\sigma^2}}\quad(e\text{는 } 2.718281\cdots\text{인 무리수})$$

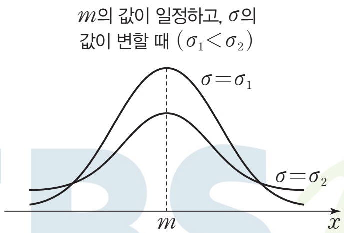
일 때 $X$의 확률분포를 정규분포라고 한다.
이때 확률변수 $X$의 평균과 표준편차는 각각 $m$, $\sigma$임이 알려져 있다.
또한 평균이 $m$, 표준편차가 $\sigma$인 정규분포를 $\mathrm{N}(m,\ \sigma^2)$과 같이 나타내고, 확률변수 $X$는 정규분포 $\mathrm{N}(m,\ \sigma^2)$을 따른다고 한다.

## 4. 정규분포를 따르는 확률변수의 확률밀도함수의 그래프

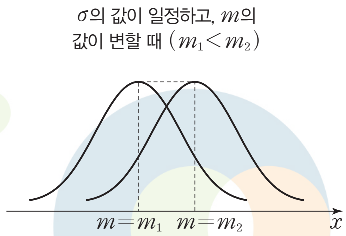



정규분포 $\mathrm{N}(m,\ \sigma^2)$을 따르는 연속확률변수 $X$의 확률밀도함수 $f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-m)^2}{2\sigma^2}}$의 그래프는 오른쪽 그림과 같은 모양이고, 다음과 같은 성질을 가지고 있음이 알려져 있다.

① 직선 $x=m$에 대하여 좌우 대칭인 종 모양의 곡선이다.
② $x$축을 점근선으로 하며, $x=m$일 때 최댓값 $\frac{1}{\sqrt{2\pi}\sigma}$을 갖는다.
③ 곡선과 $x$축 사이의 넓이는 1이다.
④ 평균 $m$의 값이 일정할 때, [그림 1]과 같이 $\sigma$의 값이 커지면 곡선의 중앙 부분이 낮아지면서 양쪽으로 퍼지고, $\sigma$의 값이 작아지면 곡선의 중앙 부분이 높아지면서 좁아지지만 대칭축의 위치는 같다.
⑤ 표준편차 $\sigma$의 값이 일정할 때, [그림 2]와 같이 $m$의 값이 변하면 대칭축의 위치는 바뀌지만 곡선의 모양은 같다.

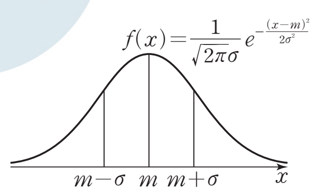


[그림 1]



[그림 2]

**참고** 확률변수 $X$가 정규분포 $\mathrm{N}(m,\ \sigma^2)$을 따를 때, 다음을 만족시킨다.
① $\mathrm{P}(X\le m)=\mathrm{P}(X\ge m)=0.5$
② $\mathrm{P}(m-\sigma\le X\le m)=\mathrm{P}(m\le X\le m+\sigma)$
③ $\mathrm{P}(m-k\sigma\le X\le m)=\mathrm{P}(m\le X\le m+k\sigma)$ (단, $k$는 양의 상수)



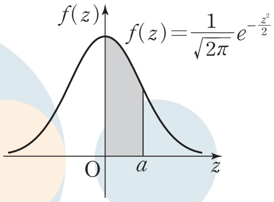
## 5. 표준정규분포

(1) 정규분포 중에서 평균이 0, 표준편차가 1인 정규분포 N(0, 1)을 표준정규분포라고 한다.

(2) 확률변수 $Z$가 표준정규분포 N(0, 1)을 따를 때, $Z$의 확률밀도함수 $f(z)$는

$$f(z)=\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}} \quad (e\text{는 } 2.718281\cdots\text{인 무리수})$$

이다. 이때 임의의 양수 $a$에 대하여 $P(0\le Z\le a)$는 오른쪽 그림에서 색칠된 부분의 넓이와 같다.




참고 확률변수 $Z$가 표준정규분포 N(0, 1)을 따를 때, $P(0\le Z\le a)$의 값은 표준정규분포표를 이용하여 구할 수 있다.
예를 들어 $P(0\le Z\le 1.96)$의 값은 표준정규분포표의 왼쪽에 있는 수 중에서 1.9를 찾고, 표의 위쪽에 있는 수 중에서 0.06을 찾아 1.9의 가로줄과 0.06의 세로줄이 만나는 곳의 수를 찾으면 된다.
즉, $P(0\le Z\le 1.96)=0.4750$이다.

| $z$ | 0.00 | 0.01 | ⋯ | 0.06 | ⋯ |
|---|---|---|---|---|---|
| 0.0 | .0000 | .0040 | ⋯ | .0239 | ⋯ |
| 0.1 | .0398 | .0438 | ⋯ | .0636 | ⋯ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 1.9 | .4713 | .4719 | ⋯ | .4750 | ⋯ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

## 6. 정규분포와 표준정규분포의 관계

확률변수 $X$가 정규분포 $N(m, \sigma^2)$ $(\sigma>0)$을 따를 때, 확률변수

$$Z=\frac{X-m}{\sigma}$$

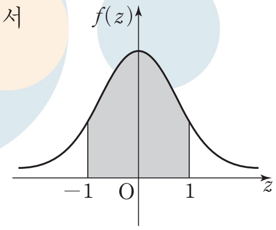
은 표준정규분포 N(0, 1)을 따른다는 사실이 알려져 있다.

이때 $P(a\le X\le b)$는 $Z=\frac{X-m}{\sigma}$을 이용하여 다음과 같이 표준정규분포 N(0, 1)을 따르는 확률변수 $Z$로 바꾸어 구한다.

$$\begin{aligned}P(a\le X\le b)&=P\left(\frac{a-m}{\sigma}\le\frac{X-m}{\sigma}\le\frac{b-m}{\sigma}\right)\\&=P\left(\frac{a-m}{\sigma}\le Z\le\frac{b-m}{\sigma}\right)\end{aligned}$$

예 확률변수 $X$가 정규분포 $N(8, 2^2)$을 따를 때, $P(6\le X\le 10)$의 값을 구해 보자.
$Z=\frac{X-8}{2}$로 놓으면 확률변수 $Z$는 표준정규분포 N(0, 1)을 따르고, 표준정규분포표에서
$P(0\le Z\le 1)=0.3413$이므로

$$\begin{aligned}P(6\le X\le 10)&=P\left(\frac{6-8}{2}\le\frac{X-8}{2}\le\frac{10-8}{2}\right)\\&=P(-1\le Z\le 1)\\&=2P(0\le Z\le 1)\\&=2\times 0.3413\\&=0.6826\end{aligned}$$



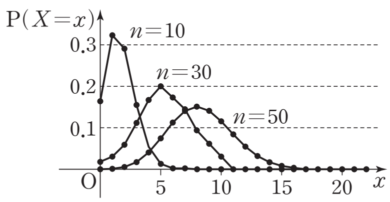
## 7. 이항분포와 정규분포의 관계

(1) 이항분포와 정규분포의 관계를 나타내는 그래프

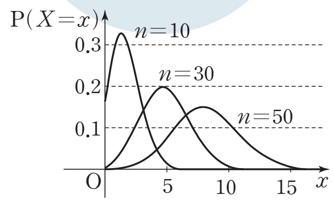
한 개의 주사위를 $n$번 던질 때, 1의 눈이 나오는 횟수를 $X$라 하면 확률변수 $X$는 이항분포 $\mathrm{B}\left(n,\ \frac{1}{6}\right)$을 따른다.

[그림 1]은 주사위를 던지는 횟수가 $n=10$, $n=30$, $n=50$일 때의 이항분포를 그래프로 나타낸 것이고, 점들을 부드럽게 연결하면 [그림 2]를 얻을 수 있다.

일반적으로 이항분포 $\mathrm{B}(n,\ p)$의 그래프는 $n$의 값이 커지면 정규분포의 확률밀도함수의 그래프에 가까워짐이 알려져 있다.




[그림 1]




[그림 2]

(2) 이항분포와 정규분포의 관계

확률변수 $X$가 이항분포 $\mathrm{B}(n,\ p)$를 따를 때, $n$이 충분히 크면 $X$는 근사적으로 정규분포 $\mathrm{N}(np,\ npq)$를 따른다. (단, $q=1-p$)

이때 확률변수 $Z=\dfrac{X-np}{\sqrt{npq}}$는 표준정규분포 $\mathrm{N}(0,\ 1)$을 따른다.

**참고** 일반적으로 $np\ge5$, $nq\ge5$이면 $n$이 충분히 큰 것으로 생각한다.

**예** 확률변수 $X$가 이항분포 $\mathrm{B}\left(100,\ \frac{1}{2}\right)$을 따를 때, $\mathrm{P}(55\le X\le60)$의 값을 구해 보자.

$$\mathrm{E}(X)=100\times\frac{1}{2}=50$$

$$\mathrm{V}(X)=100\times\frac{1}{2}\times\frac{1}{2}=25$$

이때 100은 충분히 큰 수이므로 확률변수 $X$는 근사적으로 정규분포 $\mathrm{N}(50,\ 5^2)$을 따르고, $Z=\dfrac{X-50}{5}$으로 놓으면 확률변수 $Z$는 표준정규분포 $\mathrm{N}(0,\ 1)$을 따른다.

표준정규분포표에서 $\mathrm{P}(0\le Z\le1)=0.3413$, $\mathrm{P}(0\le Z\le2)=0.4772$이므로

$$\begin{aligned}\mathrm{P}(55\le X\le60)&=\mathrm{P}\left(\frac{55-50}{5}\le\frac{X-50}{5}\le\frac{60-50}{5}\right)\\&=\mathrm{P}(1\le Z\le2)\\&=\mathrm{P}(0\le Z\le2)-\mathrm{P}(0\le Z\le1)\\&=0.4772-0.3413\\&=0.1359\end{aligned}$$

# 07 통계적 추정

## 1. 모집단과 표본

(1) 통계 조사에서 조사의 대상이 되는 집단 전체를 모집단이라 하고, 조사하기 위하여 모집단에서 뽑은 일부분을 표본이라고 한다. 이때 모집단에서 표본을 뽑는 것을 추출이라고 한다.

(2) 통계 조사에서 모집단 전체를 조사하는 것을 전수조사라 하고, 모집단의 일부분, 즉 표본을 조사하는 것을 표본조사라고 한다. 이때 표본에 포함된 대상의 개수를 표본의 크기라고 한다.

(3) 모집단에서 표본을 추출할 때, 모집단에 속하는 각 대상이 같은 확률로 추출되도록 하는 방법을 임의추출이라고 한다.

## 2. 모평균과 표본평균

(1) 어떤 모집단에서 조사하고자 하는 특성을 나타내는 확률변수를 $X$라 할 때, $X$의 평균, 분산, 표준편차를 각각 모평균, 모분산, 모표준편차라 하고, 기호로 각각 $m$, $\sigma^2$, $\sigma$와 같이 나타낸다.

(2) 모집단에서 임의추출한 크기가 $n$인 표본을 $X_1$, $X_2$, $X_3$, $\cdots$, $X_n$이라 할 때, 이 표본의 평균, 분산, 표준편차를 각각 표본평균, 표본분산, 표본표준편차라 하고, 기호로 $\overline{X}$, $S^2$, $S$와 같이 나타낸다. 이때 $\overline{X}$, $S^2$, $S$는 다음과 같이 구한다.

① $\overline{X}=\dfrac{1}{n}(X_1+X_2+X_3+\cdots+X_n)$

② $S^2=\dfrac{1}{n-1}\{(X_1-\overline{X})^2+(X_2-\overline{X})^2+(X_3-\overline{X})^2+\cdots+(X_n-\overline{X})^2\}$ (단, $n\geq 2$)

③ $S=\sqrt{S^2}$

## 3. 표본평균의 확률분포

모평균 $m$은 고정된 상수이지만 표본평균 $\overline{X}$는 임의추출된 표본에 따라 여러 가지 값을 가질 수 있으므로 확률변수이다. 따라서 $\overline{X}$의 확률분포를 구할 수 있다.

예 모집단의 확률변수 $X$의 확률분포가 오른쪽 표와 같을 때, 이 모집단에서 임의추출한 크기가 2인 표본 $(X_1, X_2)$와 그 표본평균 $\overline{X}=\dfrac{X_1+X_2}{2}$는 다음과 같다.

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

| $(X_1, X_2)$ | $(1, 1)$ | $(1, 2)$ | $(1, 3)$ | $(2, 1)$ | $(2, 2)$ | $(2, 3)$ | $(3, 1)$ | $(3, 2)$ | $(3, 3)$ |
|---|---|---|---|---|---|---|---|---|---|
| $\overline{X}$ | 1 | 1.5 | 2 | 1.5 | 2 | 2.5 | 2 | 2.5 | 3 |

따라서 확률변수 $\overline{X}$의 확률분포를 표로 나타내면 오른쪽과 같다. 이때

$$\mathrm{P}(\overline{X}=1)=\mathrm{P}(X=1)\times\mathrm{P}(X=1)=\frac{1}{3}\times\frac{1}{3}=\frac{1}{9}$$

| $\overline{X}$ | 1 | 1.5 | 2 | 2.5 | 3 | 합계 |
|---|---|---|---|---|---|---|
| $\mathrm{P}(\overline{X}=\overline{x})$ | $\frac{1}{9}$ | $\frac{2}{9}$ | $\frac{1}{3}$ | $\frac{2}{9}$ | $\frac{1}{9}$ | 1 |

## 4. 표본평균의 평균, 분산, 표준편차

모평균이 $m$, 모표준편차가 $\sigma$인 모집단에서 크기가 $n$인 표본을 임의추출할 때, 표본평균 $\overline{X}$에 대하여 다음이 성립한다.

(1) $\mathrm{E}(\overline{X})=m$

(2) $\mathrm{V}(\overline{X})=\dfrac{\sigma^2}{n}$

(3) $\sigma(\overline{X})=\dfrac{\sigma}{\sqrt{n}}$

**설명** 모집단의 확률변수 $X$의 확률분포가 오른쪽 표와 같을 때, 모평균 $m$, 모분산 $\sigma^2$, 모표준편차 $\sigma$는 각각 다음과 같다.

| $X$ | 2 | 4 | 6 | 합계 |
|---|---|---|---|---|
| $\mathrm{P}(X=x)$ | $\frac{1}{3}$ | $\frac{1}{3}$ | $\frac{1}{3}$ | 1 |

$$m=2\times\frac{1}{3}+4\times\frac{1}{3}+6\times\frac{1}{3}=4$$

$$\sigma^2=2^2\times\frac{1}{3}+4^2\times\frac{1}{3}+6^2\times\frac{1}{3}-4^2=\frac{8}{3}$$

$$\sigma=\sqrt{\frac{8}{3}}=\frac{2\sqrt{6}}{3}$$

이 모집단에서 임의추출한 크기가 2인 표본을 $(X_1,\ X_2)$라 할 때, 그 표본평균 $\overline{X}=\dfrac{X_1+X_2}{2}$가 갖는 값은 2, 3, 4, 5, 6이고 확률변수 $\overline{X}$의 확률분포를 표로 나타내면 다음과 같다.

| $\overline{X}$ | 2 | 3 | 4 | 5 | 6 | 합계 |
|---|---|---|---|---|---|---|
| $\mathrm{P}(\overline{X}=\overline{x})$ | $\frac{1}{9}$ | $\frac{2}{9}$ | $\frac{1}{3}$ | $\frac{2}{9}$ | $\frac{1}{9}$ | 1 |

$$\mathrm{E}(\overline{X})=2\times\frac{1}{9}+3\times\frac{2}{9}+4\times\frac{1}{3}+5\times\frac{2}{9}+6\times\frac{1}{9}=4$$

$$\mathrm{V}(\overline{X})=2^2\times\frac{1}{9}+3^2\times\frac{2}{9}+4^2\times\frac{1}{3}+5^2\times\frac{2}{9}+6^2\times\frac{1}{9}-4^2=\frac{4}{3}$$

$$\sigma(\overline{X})=\sqrt{\frac{4}{3}}=\frac{2\sqrt{3}}{3}$$

이때 표본의 크기가 $n=2$이므로

$$\frac{\sigma^2}{n}=\frac{8}{3}\times\frac{1}{2}=\frac{4}{3},\ \frac{\sigma}{\sqrt{n}}=\frac{2\sqrt{6}}{3}\times\frac{1}{\sqrt{2}}=\frac{2\sqrt{3}}{3}$$

따라서 다음이 성립한다.

① $\mathrm{E}(\overline{X})=4=m$

② $\mathrm{V}(\overline{X})=\dfrac{4}{3}=\dfrac{\sigma^2}{n}$

③ $\sigma(\overline{X})=\dfrac{2\sqrt{3}}{3}=\dfrac{\sigma}{\sqrt{n}}$

## 5. 표본평균의 분포

모평균이 $m$, 모표준편차가 $\sigma$인 모집단에서 임의추출한 크기가 $n$인 표본의 표본평균 $\overline{X}$에 대하여 다음이 성립한다.

(1) 모집단이 정규분포 $\mathrm{N}(m, \sigma^2)$을 따르면 표본평균 $\overline{X}$는 정규분포 $\mathrm{N}\left(m, \frac{\sigma^2}{n}\right)$을 따른다.

(2) 모집단이 정규분포를 따르지 않을 때에도 표본의 크기 $n$이 충분히 크면 표본평균 $\overline{X}$는 근사적으로 정규분포 $\mathrm{N}\left(m, \frac{\sigma^2}{n}\right)$을 따른다.

예 ① 모집단의 확률변수 $X$가 정규분포 $\mathrm{N}(50, 8^2)$을 따르고, 이 모집단에서 크기가 4인 표본을 임의추출할 때 표본평균을 $\overline{X}$라 하면

$$m=\mathrm{E}(X)=50$$
$$\sigma^2=\mathrm{V}(X)=8^2=64$$

이므로

$$\mathrm{E}(\overline{X})=\mathrm{E}(X)=50$$
$$\mathrm{V}(\overline{X})=\frac{\mathrm{V}(X)}{n}=\frac{64}{4}=16$$

따라서 표본평균 $\overline{X}$는 정규분포 $\mathrm{N}(50, 4^2)$을 따른다.

② 정규분포 $\mathrm{N}(50, 8^2)$을 따르는 모집단에서 크기가 4인 표본을 임의추출할 때 표본평균 $\overline{X}$는 정규분포 $\mathrm{N}(50, 4^2)$을 따르므로 확률변수 $Z=\frac{\overline{X}-50}{4}$은 표준정규분포 $\mathrm{N}(0, 1)$을 따른다. 따라서 $\mathrm{P}(\overline{X}\geq 54)$의 값을 오른쪽 표준정규분포표를 이용하여 구해 보면 다음과 같다.

| $z$ | $\mathrm{P}(0\leq Z\leq z)$ |
|---|---|
| 0.5 | 0.1915 |
| 1.0 | 0.3413 |
| 1.5 | 0.4332 |
| 2.0 | 0.4772 |

$$\begin{aligned}\mathrm{P}(\overline{X}\geq 54)&=\mathrm{P}\left(\frac{\overline{X}-50}{4}\geq\frac{54-50}{4}\right)\\&=\mathrm{P}(Z\geq 1)\\&=0.5-\mathrm{P}(0\leq Z\leq 1)\\&=0.5-0.3413\\&=0.1587\end{aligned}$$

## 6. 모평균의 추정

(1) 모집단에서 추출한 표본에서 얻은 자료를 이용하여 모집단의 어떤 성질을 확률적으로 추측하는 것을 추정이라고 한다.

(2) 정규분포 $\mathrm{N}(m, \sigma^2)$을 따르는 모집단에서 크기가 $n$인 표본을 임의추출하여 구한 표본평균 $\overline{X}$의 값이 $\overline{x}$일 때, 모평균 $m$에 대한 신뢰구간은 다음과 같다.

① 신뢰도 95 %의 신뢰구간

$$\overline{x}-1.96\frac{\sigma}{\sqrt{n}}\le m\le \overline{x}+1.96\frac{\sigma}{\sqrt{n}}$$

② 신뢰도 99 %의 신뢰구간

$$\overline{x}-2.58\frac{\sigma}{\sqrt{n}}\le m\le \overline{x}+2.58\frac{\sigma}{\sqrt{n}}$$

**설명** ① 모집단의 분포가 정규분포 $\mathrm{N}(m, \sigma^2)$을 따를 때, 크기가 $n$인 표본을 임의추출하면 표본평균 $\overline{X}$는 정규분포 $\mathrm{N}\left(m, \dfrac{\sigma^2}{n}\right)$을 따르고, 확률변수 $Z=\dfrac{\overline{X}-m}{\frac{\sigma}{\sqrt{n}}}$은 표준정규분포 $\mathrm{N}(0, 1)$을 따른다.

이때 표준정규분포표에서 $\mathrm{P}(-1.96\le Z\le 1.96)=0.95$이므로

$$\begin{aligned}\mathrm{P}\left(-1.96\le \frac{\overline{X}-m}{\frac{\sigma}{\sqrt{n}}}\le 1.96\right)&=\mathrm{P}\left(-1.96\frac{\sigma}{\sqrt{n}}\le \overline{X}-m\le 1.96\frac{\sigma}{\sqrt{n}}\right)\\&=\mathrm{P}\left(-1.96\frac{\sigma}{\sqrt{n}}\le m-\overline{X}\le 1.96\frac{\sigma}{\sqrt{n}}\right)\\&=\mathrm{P}\left(\overline{X}-1.96\frac{\sigma}{\sqrt{n}}\le m\le \overline{X}+1.96\frac{\sigma}{\sqrt{n}}\right)\\&=0.95\end{aligned}$$

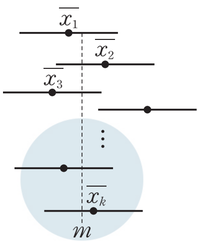
따라서 모집단으로부터 크기가 $n$인 표본을 임의추출하여 구한 표본평균 $\overline{X}$의 값을 $\overline{x}$라 할 때,

$$\overline{x}-1.96\frac{\sigma}{\sqrt{n}}\le m\le \overline{x}+1.96\frac{\sigma}{\sqrt{n}}$$

를 모평균 $m$에 대한 신뢰도 95 %의 신뢰구간이라고 한다.

② 표본평균 $\overline{X}$는 확률변수이므로 추출되는 표본에 따라 표본평균 $\overline{X}$의 값 $\overline{x}$가 달라지고 그에 따라 신뢰구간도 달라진다. 이와 같은 신뢰구간 중에는 오른쪽 그림과 같이 모평균 $m$을 포함하는 것과 포함하지 않는 것이 있을 수 있다.

따라서 모평균 $m$에 대하여 신뢰도 95 %의 신뢰구간이란 크기가 $n$인 표본을 여러 번 임의추출하여 신뢰구간을 각각 구하면 그 중에서 95 %는 모평균 $m$을 포함할 것으로 기대된다는 것을 의미한다.




**참고** 모평균의 신뢰구간을 구할 때 모표준편차 $\sigma$의 값을 알 수 없는 경우가 많다. 이 경우 표본의 크기 $n$이 충분히 크면 모표준편차 $\sigma$ 대신 표본표준편차 $s$를 이용하여 모평균의 신뢰구간을 구할 수 있다는 것이 알려져 있다.