

AP Statistics

Sungchan Yi

January 2019

Contents

1	순열과 조합	2
2	확률의 뜻과 활용	5
3	조건부확률	7
4	통계학	9

1 순열과 조합

정의 1.1. $0! = 1$, $n! = \prod_{i=1}^n i = n \cdot (n-1) \cdots 2 \cdot 1$ ($n \geq 1$) 로 정의하고, $!$ 는 **팩토리얼**(factorial) 이라 읽는다.

정의 1.2. 서로 다른 n 개의 원소에서 서로 다른 r 개를 택하여 일렬로 배열하는 것을 n 개에서 r 개를 택하는 **순열**(permutation)이라 하고, 기호로 ${}_nP_r$ 와 같이 나타낸다.

정리 1.3. ${}_nP_r = n(n-1) \cdots (n-r+1) = \frac{n!}{(n-r)!}$ (단, $0 \leq r \leq n$)

정의 1.4. 서로 다른 n 개의 원소에서 순서를 생각하지 않고 r 개를 택하는 것을 n 개에서 r 개를 택하는 **조합**(combination)이라 하고, 기호로 ${}_nC_r$ 또는 $\binom{n}{r}$ 과 같이 나타낸다.

정리 1.5. $\binom{n}{r} = \frac{{}_nP_r}{r!} = \frac{n!}{r!(n-r)!}$ (단, $0 \leq r \leq n$)

정리 1.6. (조합의 성질)

$$(1) \quad \binom{n}{r} = \binom{n}{n-r} \quad (\text{단, } 0 \leq r \leq n) \quad (\text{대칭성})$$

$$(2) \quad \binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1} \quad (\text{단, } 1 \leq r \leq n-1) \quad (\text{파스칼 법칙})$$

정의 1.7. 서로 다른 n 개의 원소에서 중복을 허락하여 r 개를 택하는 순열을 n 개에서 r 개를 택하는 **중복순열**이라 하고, 기호로 ${}_n\Pi_r$ 과 같이 나타낸다.

정의 1.8. 서로 다른 n 개의 원소에서 중복을 허락하여 r 개를 택하는 조합을 n 개에서 r 개를 택하는 **중복조합**이라 하고, 기호로 ${}_nH_r$ 과 같이 나타낸다.

정리 1.9. ${}_n\Pi_r = n^r$, ${}_nH_r = \binom{n+r-1}{r}$.

정리 1.10. $n \in \mathbb{N}$ 에 대하여,

$$(x+y)^n = \sum_{r=0}^n \binom{n}{r} x^{n-r} y^r = \binom{n}{0} x^n y^0 + \binom{n}{1} x^{n-1} y^1 + \cdots + \binom{n}{n} x^0 y^n$$

이다. 이를 $(a+b)^n$ 에 대한 **이항정리**(binomial theorem)라 하고, $\binom{n}{r} x^{n-r} y^r$ 을 전개식의 **일반항**, 전개식의 각 항의 계수 $\binom{n}{r}$ 들을 **이항계수**라 한다.

정리 1.11. (이항계수의 성질)

$$(1) \quad (1+x)^n = \sum_{r=0}^n \binom{n}{r} x^r = \binom{n}{0} + \binom{n}{1} x + \cdots + \binom{n}{n} x^n \quad (\text{for all } x \in \mathbb{C})$$

$$(2) \quad \sum_{r=0}^n \binom{n}{r} = \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n} = 2^n$$

$$(3) \quad \sum_{r=0}^n (-1)^r \binom{n}{r} = \binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \binom{n}{3} + \cdots + (-1)^n \binom{n}{n} = 0$$

$$(4) \quad \sum_{r=0}^n r \binom{n}{r} = \binom{n}{1} + 2 \cdot \binom{n}{2} + \cdots + n \cdot \binom{n}{n} = n \cdot 2^{n-1}$$

$$(5) \quad \sum_{r=0}^n r^2 \binom{n}{r} = 2^2 \cdot \binom{n}{2} + 3^2 \cdot \binom{n}{3} + \cdots + n^2 \cdot \binom{n}{n} = n(n+1) \cdot 2^{n-2}$$

$$(6) \quad \sum_{r=0}^n \frac{1}{r+1} \binom{n}{r} = \frac{1}{1} \binom{n}{0} + \frac{1}{2} \binom{n}{1} + \cdots + \frac{1}{n+1} \binom{n}{n} = \frac{1}{n+1} (2^{n+1} - 1)$$

정리 1.12. $n \in \mathbb{N}$ 에 대하여,

$$(x_1 + x_2 + \cdots + x_m)^n = \sum_{r_1+r_2+\cdots+r_m=n} \binom{n}{r_1, r_2, \dots, r_m} x_1^{r_1} x_2^{r_2} \cdots x_m^{r_m}$$

이고, 이를 $(x_1 + x_2 + \cdots + x_m)^n$ 에 대한 **다항정리**(multinomial theorem)라 한다. 이 때 $\binom{n}{r_1, r_2, \dots, r_m}$ 를 **다항계수**라 하고, 다음과 같이 정의한다.

$$\binom{n}{r_1, r_2, \dots, r_m} = \frac{n!}{r_1! \cdot r_2! \cdot \cdots \cdot r_m!}$$

정의 1.13. 서로 다른 n 개의 원소를 원형으로 배열하는 순열을 **원순열**이라 하고, 그 경우의 수는 $(n-1)!$ 이다.

정리 1.14. (원순열의 일반공식) n 개 중에서 서로 같은 것이 p_1, p_2, \dots, p_k 개씩 있을 때, 이 $n (= p_1 + \dots + p_k)$ 개를 원형으로 배열하는 방법(원순열)의 수는 다음과 같다.

$$\frac{1}{n} \sum_{d|g} \left\{ \phi(d) \binom{\frac{n}{d}}{\frac{p_1}{n}, \frac{p_2}{n}, \dots, \frac{p_k}{n}} \right\}$$

단, $g = \gcd(p_1, \dots, p_k)$, $d > 0$ 이고 $\phi(d)$ 는 d 이하의 자연수 중에서 d 와 서로소인 자연수의 개수로 정의된다.

2 확률의 뜻과 활용

정의 2.1. 같은 조건 아래에서 반복할 수 있고, 그 결과가 우연에 의하여 결정되는 실험이나 관찰을 **시행**이라고 한다. 어떤 시행에서 일어날 수 있는 모든 가능한 결과 전체의 집합을 **표본공간**(sample space S)이라 하고, 표본공간의 부분집합을 **사건**(event)이라고 한다.

정의 2.2. 표본공간의 부분집합 중에서 원소의 개수가 한 개인 집합을 **근원사건**이라 하고, 반드시 일어나는 사건은 **전사건**, 절대로 일어나지 않는 사건은 **공사건**(\emptyset)이라 한다.

정의 2.3. 두 사건 A, B 에 대하여, A 또는 B 가 일어나는 사건을 A 와 B 의 **합사건**이라 하고, $A \cup B$ 로 나타낸다. 그리고 A 와 B 가 동시에 일어나는 사건을 A 와 B 의 **곱사건**이라 하고, $A \cap B$ 로 나타낸다.

정의 2.4. 표본공간 S 의 부분집합인 두 사건 A, B 에 대하여 $A \cap B = \emptyset$ 이면 A 와 B 는 서로 **배반사건**이라 한다. 또, 사건 A 가 일어나지 않는 사건을 사건 A 의 **여사건**이라 하고, A^C 로 나타낸다.

정의 2.5. 어떤 시행에서 사건 A 가 일어날 가능성을 수로 나타낸 것을 사건 A 가 일어날 **확률**이라 하고, 기호로 $P(A)$ 와 같이 나타낸다.

정의 2.6. (수학적 확률) 어떤 시행의 표본공간 S 가 m 개의 근원사건으로 이루어져 있고, 각 근원사건이 일어날 가능성이 모두 같은 정도로 기대될 때, 사건 A 가 r 개의 근원사건으로 이루어져 있으면 사건 A 가 일어날 확률은 다음과 같다.

$$P(A) = \frac{(\text{사건 } A \text{가 일어나는 경우의 수})}{(\text{모든 경우의 수})} = \frac{n(A)}{n(S)} = \frac{r}{m}$$

정의 2.7. (통계적 확률) 같은 시행을 n 번 반복하여 사건 A 가 일어난 횟수를 r_n 이라고 하자. 이 때, 시행 횟수 n 이 한없이 커짐에 따라 그 상대도수 r_n/n 은 $P(A)$ 에 가까워진다.

$$P(A) = \lim_{n \rightarrow \infty} \frac{r_n}{n}$$

정의 2.8. (기하학적 확률) 연속적인 변량을 크기로 갖는 표본공간의 영역 S 안에서 각각의 점을 잡을 가능성이 같은 정도로 기대될 때, 영역 S 에 포함되어 있는 영역 A 에 대하여 영역 S 에서 임의로 잡은 점이 영역 A 에 속할 확률은 다음과 같다.

$$P(A) = \frac{(\text{영역 } A \text{의 크기})}{(\text{영역 } S \text{의 크기})}$$

정의 2.9. (확률의 공리 - Axioms of Probability) 표본공간 S 와 사건 A 에 대하여,

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

(2) $P(S) = 1$

(3) 서로 배반인 사건열 A_1, A_2, \dots 에 대해 $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$

정리 2.10. (확률의 기본 성질)

(1) $P(\emptyset) = 0$

(2) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (확률의 덧셈정리)

(3) $P(A^C) = 1 - P(A)$ (여사건의 확률)

정리 2.11. (포함 배제 원리) 사건 A_1, \dots, A_n 에 대하여 다음이 성립한다.

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n (-1)^{k+1} \left(\sum_{1 \leq i_1 < \dots < i_k \leq n} P(A_{i_1} \cap \dots \cap A_{i_k}) \right)$$

3 조건부확률

정의 3.1. 확률이 0이 아닌 두 사건 A, B 에 대하여 사건 A 가 일어났을 때, 사건 B 가 일어날 확률을 사건 A 가 일어났을 때의 사건 B 의 **조건부확률**(conditional probability)이라 하고, 기호로 $P(B|A)$ 와 같이 나타낸다. 이는 다음과 같이 계산한다.

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \quad (P(A) > 0)$$

정리 3.2. (확률의 곱셈정리) 공사건이 아닌 두 사건 A, B 에 대하여 다음이 성립한다.

$$P(A \cap B) = P(B|A)P(A) = P(A|B)P(B)$$

정의 3.3. 두 사건 A, B 에 대하여 사건 A 가 일어났을 때의 사건 B 의 조건부확률이 사건 B 가 일어날 확률과 같을 때, 즉

$$P(B|A) = P(B|A^C) = P(B)$$

이면, 두 사건 A, B 는 서로 **독립**(independent)이라 하고, 기호로 $A \perp B$ 와 같이 나타낸다. 두 사건이 독립이 아닐 때는 **종속**이라 한다.

정리 3.4. 공사건이 아닌 두 사건 A, B 에 대하여 다음 조건은 서로 동치이다.

- (1) A, B 가 서로 독립이다.
- (2) $P(A \cap B) = P(A)P(B)$

정리 3.5. 공사건이 아닌 두 사건 A, B 에 대하여 다음이 성립한다.

$$[A, B \text{가 독립}] \iff [A^C, B \text{가 독립}] \iff [A, B^C \text{가 독립}] \iff [A^C, B^C \text{가 독립}]$$

정의 3.6. 공사건이 아닌 사건 A_1, \dots, A_n 에 대하여,

$$P(A_i \cap A_j) = P(A_i)P(A_j) \quad \text{for all } 1 \leq i \neq j \leq n$$

이 성립하면 사건 A_1, \dots, A_n 이 **쌍마다 독립**(pairwise independent)이라고 한다.

정의 3.7. 공사건이 아닌 사건 A_1, \dots, A_n 에 대하여,

$$P\left(\bigcap_{i=1}^n A_i\right) = \prod_{i=1}^n P(A_i)$$

이 성립하면 사건 A_1, \dots, A_n 이 **상호 독립**(mutually independent)이라고 한다.

정의 3.8. 동일한 시행을 반복할 때, 각 시행에서 일어나는 사건이 서로 독립이면 이러한 시행을 **독립시행**이라고 한다.

정리 3.9. (독립시행의 확률) 1회의 시행에서 사건 A 가 일어날 확률을 p 라 할 때, 이 시행을 n 회 반복하는 독립시행에서 사건 A 가 r 번 일어날 확률은 다음과 같다.

$$\binom{n}{r} p^r (1-p)^{n-r} \quad (r = 0, 1, \dots, n)$$

4 자료의 생성