1장 확률

§1.1. 표본공간, 사건

Definition

(1) **확률**(probability) :

일정한 조건 하에서 어떤 사건이 일어날 가능성의 정도를 수치로 나타낸 것이다.

(2) **통계**(statistics) :

사회 현상부터 자연 현상에 이르기까지 다양한 자료를 일정한 체계에 따라 수치의 형태로 분석, 표현해 낸 것이다.

Definition

(1) **확률현상**(random phenomenon) :

사회 현상이나 자연 현상 중 결과가 미리 결정되지 않는 현상, 즉 결과가 어떤 불확실성에 의해서 좌우되는 현상을 확률 현상이라고 한다.

(2) **확률실험**(random experiment) :

결과가 확률현상으로 나타나는 실험이나 관찰 또는 조사를 확률실험이라 한다.

참고 확률실험의 예 : 동전 던지기,주사위 던지기, 카드 뽑기

Definition

(1) 확률실험에서 얻을 수 있는 가능한 모든 결과들의 집합을 그 확률실험의 **표본공간**(sample space)이라 하고, 기호로 S 또는 Ω로 나타낸다.

(2) 표본공간안의 원소 하나하나를 **표본점**(sample point)이라 하고 $\omega_1, \omega_2, \omega_3, \cdots$ 등으로 나타낸다.

(3) 표본공간의 부분집합으로 어떤 조건을 만족하는 특정한 표본점들의 집합을 **사건**(event)이라 하고, 대문자 A, B, C, $\cdots$등으로 표시한다.

(4) 표본공간 Ω를 **전사건**(total event), $\varnothing$를 **공사건**(empty event)이라 하고, 특히 표본점 하나로 이루어진 사건을 **근원사건**(elementary event)이라 한다.

참고 1 하나의 동전을 던지는 확률실험에서 표본공간

$$\Omega = \{ H(\text{앞면}), T(\text{뒷면}) \}$$

2 하나의 주사위 던지는 확률실험에서 표본공간

$$\Omega = \{ ⚀, ⚁, ⚂, ⚃, ⚄, ⚅ \} = \{ 1, 2, 3, 4, 5, 6 \}$$

Example

(1) 동전을 반복해서 세 번 던지는 확률실험에서 표본공간 Ω를 구하여라.

(2) 적어도 한 번 앞면(H)이 나올 사건 A를 구하여라.

(3) 앞면(H)보다 뒷면(T)이 많이 나올 사건 B를 구하여라.

Example

(1) 주사위를 반복해서 두 번 던지는 확률실험에서 나온 눈의 개수에 대한 표본공간 Ω을 구하여라.

(2) 첫 번째 나온 눈의 개수와 두 번째 나온 눈의 개수가 같은 사건 A를 구하여라.

(3) 첫 번째 나온 눈의 개수가 두 번째 나온 눈의 개수의 두 배가 되는 사건 B를 구하여라.

Example

1에서 5까지의 숫자가 적힌 공이 들어 있는 주머니에서 반복하여 2개의 공을 꺼낸다고 할 때,

(1) 처음에 꺼낸 공을 주머니에 다시 넣고 두 번째 공을 꺼내는 확률실험의 표본공간 Ω_1을 구하여라.

(2) 처음에 꺼낸 공을 주머니에 다시 넣지 않고 두 번째 공을 꺼내는 확률실험의 표본공간 Ω_2을 구하여라.

Definition

확률실험에서 같은 시행을 두 번 이상 반복 할 때,

(1) 추출하였던 것을 제자리에 되돌려 놓고 다음 것을 추출하는 방법을 **복원추출**(replacement)이라 하고,

(2) 추출하였던 것을 제자리에 되돌리지 않고 다음 것을 추출하는 방법을 **비복원추출**(without replacement)이라 한다.

Definition

표본공간 Ω의 부분집합인 두 사건 A, B에 대하여

(1) A와 B의 **합사건**(union of events) :

사건 A 또는 B가 일어나는 사건으로

$$A \cup B = \{\omega \in \Omega \mid \omega \in A \text{ 또는 } \omega \in B\}$$

로 나타낸다.

(2) A와 B의 **곱사건**(intersection of events) :

사건 A와 B가 동시에 일어나는 사건으로

$$A \cap B = \{\omega \in \Omega \mid \omega \in A \text{이고 } \omega \in B\}$$

로 나타낸다.

참고 유한곱사건 : $\bigcap_{i=1}^{n} A_i = \{\omega \in \Omega \mid \omega \in A_i \text{ for all } 1 \le i \le n\}$

유한합사건 : $\bigcup_{i=1}^{n} A_i = \{\omega \in \Omega \mid \omega \in A_i \text{ for some } 1 \le i \le n\}$

무한가산곱사건 : $\bigcap_{i=1}^{\infty} A_i = \{\omega \in \Omega \mid \omega \in A_i \text{ for all } 1 \le i < \infty\}$

무한가산합사건 : $\bigcup_{i=1}^{\infty} A_i = \{\omega \in \Omega \mid \omega \in A_i \text{ for some } 1 \le i < \infty\}$

Definition

표본공간 Ω의 부분집합인 두 사건 A, B에 대하여

(1) A의 **여사건**(complementary event) :

사건 A가 일어나지 않는 사건으로

$$A^C = \{\omega \in \Omega \mid \omega \notin A\}$$

로 나타낸다.

(2) A와 B의 **차사건**(difference of events) :

사건 A에는 있으나 사건 B에는 없는 사건으로

$$\begin{aligned} A - B &= A \cap B^C \\ &= \{\omega \in \Omega \mid \omega \in A \text{ 그리고 } \omega \notin B\} \end{aligned}$$

로 나타낸다.

Definition

(1) 표본공간 Ω의 부분집합인 두 사건 A와 B가 동시에 일어나지 않을 때, 즉 $A \cap B = \varnothing$이면, 두 사건 A, B를 **배반사건**(mutually exclusive events)이라 한다.

(2) 표본공간 Ω의 부분집합인 n개의 사건들 $A_1, A_2, \cdots, A_n$에 대하여

$$A_i \cap A_j = \varnothing \quad (i \ne j,\ 1 \le i, j \le n)$$

일 때, $A_1, A_2, \cdots, A_n$을 **쌍마다 배반사건**(pairwisely mutually exclusive events)이라 한다.

Definition

표본공간 Ω의 부분집합인 n개의 사건들 $A_1, A_2, \cdots, A_n$에 대하여 다음 두 조건

① $A_1, A_2, \cdots, A_n$은 쌍마다 배반사건이다.
즉, $A_i \cap A_j = \varnothing\ (1 \le i, j \le n$ and $i \ne j)$이다.

② $\bigcup_{i=1}^{n} A_i = \Omega$

을 만족하면, $\{A_i \mid i = 1, 2, \cdots, n\}$을 표본공간 Ω의 **분할**(partition)이라고 한다.

Example

공정한 동전을 세 번 던지는 게임에서 앞면(H)가 나온 횟수가 i인 사건을 $A_i\ (i = 0, 1, 2, 3)$라 하면, 이 사건들은 표본공간 Ω의 분할임을 보여라.

Example

공정한 동전을 세 번 던지는 게임에서 앞면(H)이 두 번 이상 나온 사건을 A, 뒷면(T)이 한 번 이상 나오는 사건을 B라 할 때,

$$A \cup B,\ A \cap B$$

를 구하고, 두 사건 A와 B는 표본공간 Ω의 분할이 아님을 보여라.

§1.2. 확률

Definition

한 확률실험에서

'어떤 사건 A가 일어날 가능성의 정도'

를 0과 1사이의 실수 값으로 표현한 것을 사건 A의 `**확률**(probability)이라 하고, 이 값을 기호로 $P(A)$로 나타낸다.

참고 확률의 정의로는 보통 다음 세 가지가 이용된다.

1. 수학적 확률(사전 확률)
2. 통계적 확률(사후 확률)
3. 공리적 확률

1 **수학적 확률**(mathematical probability)

Definition

한 확률실험에서 모든 결과가 동일한 가능성을 갖고 일어난다고 하자.

이 확률실험에서 표본공간 Ω안의 표본점의 개수가 N, 사건 A안의 표본점의 개수가 n이면,

$$P(A) = \frac{\text{사건 } A\text{안의 표본점의 개수}}{\text{표본공간 } \Omega\text{안의 표본점의 개수}} = \frac{n}{N}$$

를 사건 A의 **수학적 확률**이라고 한다.

단점 수학적 확률은 확률실험에서 얻은 결과의 수가 **유한**인 경우에 한하여 적용된다.

2 **통계적 확률** (empirical probability)

Definition

한 확률실험을 N번 반복한 시행에서 사건 A가 발생한 횟수를 $n(A)$이라 할 때, 사건 A의 **상대도수**(relative frequency)

$$P(A) = \frac{\text{사건 } A\text{가 발생한 횟수}}{\text{전체 시행 횟수}} = \frac{n(A)}{N}$$

를 사건 A의 **통계적 확률**이라고 한다.

단점 시행횟수에 따라 **통계적 확률**이 다르게 나오게 된다.

큰 수의 법칙 (law of large number) :

확률실험의 시행이 많아질수록 통계적 확률은 수학적 확률에 가까워진다.

$$P(A) = \lim_{N\to\infty} \frac{n(A)}{N}$$

Example

(1) 공정한 동전을 한 번 던져서 앞면(H)이 나타날 확률을 구하여라.

(2) 동전을 10000번 던지는 확률실험을 통하여 다음의 결과를 얻었다. 이 결과를 기초로 앞면(H)이 나타날 확률을 구하여라.

던진 횟수	앞면의 수	뒷면의 수
10	4	6
100	28	22
1000	514	486
10000	5017	4983

Example

공정한 동전을 세 번 던지는 게임에서 앞면(H)이 나온 횟수가 i인 사건을 A_i $(i=0,1,2,3)$ 이라 할 때, 이 사건들의 확률을 구하여라.

Example

공정한 주사위를 반복해서 두 번 던지는 게임에서 첫 번째 나온 눈의 개수가 3의 배수인 사건을 A, 두 번째 나온 눈의 개수가 3의 배수인 사건을 B라 할 때, $P(A)$와 $P(A\cap B)$를 구하여라.

▣ 확률의 기본성질

Theorem

임의의 사건 A와 B에 대하여 다음 성질이 성립한다.

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

(2) $P(A^C)=1-P(A)$

(3) $P(A\cup B)=P(A)+P(B)-P(A\cap B)$

(4) $P(A\cup B)\le P(A)+P(B)$

(5) A와 B가 배반이면, $P(A\cup B)=P(A)+P(B)$

(6) $A\subset B$이면, $P(B-A)=P(B)-P(A)$

(7) $A\subset B$이면, $P(A)\le P(B)$

참고 임의의 사건 A, B, C에 대하여 다음 성질이 성립한다.

$$\begin{aligned}P(A\cup B\cup C)&=P(A)+P(B)+P(C)\\&\quad-P(A\cap B)-P(B\cap C)-P(C\cap A)\\&\quad+P(A\cap B\cap C)\end{aligned}$$

Example

공정한 주사위를 반복해서 두 번 던지는 게임에서 첫 번째 나온 눈의 개수가 3의 배수인 사건을 A, 두 번째 나온 눈의 개수가 4의 배수인 사건을 B, 그리고 두 눈의 개수의 합이 7인 사건을 C라 할 때, 다음 확률을 구하여라.

(1) $P(A\cup C)$

(2) $P(B\cup C)$

(3) $P(A\cup B\cup C)$

3 공리론적 확률(axiomatic probability)

Definition

표본공간 Ω에서 정의된 모든 사건들의 모임을

$$\mathfrak{F}=\wp(\Omega)=\{A\mid A\subset\Omega\}$$

라 할 때, 만약 집합 $\mathfrak{F}$ 위에서 정의된 함수

$$\begin{aligned}P:\ &\mathfrak{F}\to\mathbb{R}\\&\ \Downarrow\\&A\mapsto P(A)\end{aligned}$$

가 다음 세 공리

$(A1)$ 모든 $A\in\mathfrak{F}$에 대하여, $0\le P(A)\le 1$이다.
$(A2)$ $P(\Omega)=1$
$(A3)$ 쌍마다 배반인 사건들 A_1, A_2, $\cdots$에 대하여

$$P(A_1\cup A_2\cup\cdots)=\sum_{i=1}^{\infty}P(A_i)$$

를 만족하면, P를 $\mathfrak{F}$위에서 정의된 **확률함수**(probability function) 또는 **확률측도**(probability measure)라 하며 $P(A)$를 사건 A의 **확률**(probability)이라 한다.

이때 $(\Omega,\mathfrak{F},P)$를 **확률공간**(probability space)이라 한다.

4 기하적 확률(geometric probability)

Definition

표본공간 Ω가 비가산집합일 때, 만약 표본공간 Ω에서 길이(ℓ)나 면적(area) 또는 체적(V) 등 유한의 기하학적인 측도 $m(\Omega)$가 주어진 경우, 사건 A의 확률은

$$P(A)=\frac{\ell(A)}{\ell(\Omega)}\ \text{또는}\ \frac{\text{area}(A)}{\text{area}(\Omega)}\ \text{또는}\ \frac{V(A)}{V(\Omega)}$$

와 같이 정의된다.

이때 P는 확률함수의 세 공리 $(A1)$,$(A2)$,$(A3)$을 만족한다.

Example

어떤 커플은 정오부터 1시 사이에 약속장소에서 만나기로 하였고, 누가 먼저 약속 장소에 도착하든지 10분 이상 기다리지 않기로 하였다. 이 커플이 만날 확률을 구하여라.

§1.3. 조건부 확률

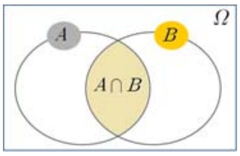
Definition

확률공간 $(\Omega,\mathfrak{F},P)$에서, $A,\ B\in\mathfrak{F}$이고 $P(A)>0$일 때, 사건 A가 일어났다는 조건아래서 사건 B가 일어날 확률을 사건 A에 대한 사건 B의 **조건부확률**(conditional probability)이라 하고, 다음과 같이 정의한다.

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

참고 집합 $\mathfrak{F}$위에서 정의된 함수

$$P(\cdot|A):\mathfrak{F}\to\mathbb{R}$$
$$B\mapsto P(B|A)$$



는 확률함수이다.

$(A1)$ 모든 $B\in\mathfrak{F}$에 대하여 $0\le P(B|A)\le 1$이다.

$(A2)$ $P(\Omega|A)=1$

$(A3)$ 쌍마다 배반인 사건들 $A_1,\ A_2,\ \cdots$에 대하여

$$P(\bigcup_{i=1}^{\infty}A_i|A)=\sum_{i=1}^{\infty}P(A_i|A)$$

즉, $(\Omega,\mathfrak{F},P(\cdot|A))$는 확률공간이다.

Example

주사위를 두 번 던지는 확률실험에서 첫 번째 나온 눈의 개수가 3의 배수 사건을 A, 두 눈의 개수의 합이 7인 사건을 B라 할 때, 사건 A에 대한 사건 B의 조건부 확률을 구하여라.

Example

다음 표와 같은 어느 대학의 신입생들 중에서 임의로 한 명을 선출한다고 한다. 이때, 다음 조건부 확률을 구하여라.

성별 \ 출신지	대도시	중소도시	농어촌	기타	계
여학생	442	276	146	4	868
남학생	1,145	662	313	12	2,132
계	1,587	938	459	16	3,000

(1) 선출된 학생이 여학생일 때, 이 학생이 농어촌 출신일 확률

(2) 선출된 학생이 남학생일 때, 이 학생이 대도시 출신일 확률

(3) 선출된 학생이 중소도시 출신일 때, 이 학생이 여학생일 확률

Example

주사위를 두 번 반복하여 던지는 실험에서 첫 번째 나온 눈의 개수가 2의 배수인 사건을 A, 두 번째 나온 눈의 개수가 5인 사건을 B라 할 때, 다음 확률을 구하여라.

(1) $P(B|A)$　　　　(2) $P(A|B)$

Theorem **곱의 법칙**(Multiplication Law)

(1) 사건 $A_1,\ A_2$에 대하여 $P(A_1)>0$일 때,

$$P(A_1\cap A_2)=P(A_1)P(A_2|A_1)$$

(2) 사건 $A_1,\ A_2,\ A_3$에 대하여 $P(A_1\cap A_2)>0$일 때,

$$P(A_1\cap A_2\cap A_3)=P(A_1)P(A_2|A_1)P(A_3|A_1\cap A_2)$$

(3) 사건 $A_1,\cdots,A_n$에 대하여 $P(A_1\cap\cdots\cap A_{n-1})>0$일 때,

$$P(A_1\cap A_2\cap\cdots\cap A_n)$$
$$=P(A_1)P(A_2|A_1)P(A_3|A_1\cap A_2)\cdots P(A_n|A_1\cap\cdots\cap A_{n-1})$$

(proof)

Example

카드 두 장을 차례로 뽑는 게임을 할 때,

(1) 비복원추출에 의해 뽑은 두 장의 카드가 모두 하트일 확률을 구하여라.

(2) 복원추출에 의해 뽑은 두 장의 카드가 차례로 하트와 다이아몬드일 확률을 구하여라.

Example

빨간 공 3개와 흰 공 5개 그리고 검은 공 2개가 들어 있는 주머니에서 비복원 추출에 의하여 무작위로 공 3개를 차례로 꺼낼 때, 흰 공과 검은 공 그리고 빨간 공의 순서로 나올 확률을 구하여라.

▣ 사건의 독립성

> ***Definition***
>
> (1) 두 사건 A, B에 대하여, $P(A)>0$일 때
>
> $$P(B)=P(B|A)$$
>
> 이면, 사건 B는 사건 A와 **독립**(independent)이라 한다.
>
> (2) 두 사건 A, B에 대하여,
>
> $$P(A\cap B)=P(A)P(B)$$
>
> 이면, A와 B는 **서로 독립**(independent)이라 한다.
> 만약 A와 B가 서로 독립이 아니면, A와 B는 **서로 종속**(dependent)이라고 한다.

참고 $\boxed{1}$ A와 B가 서로 독립이면, A와 B^C도 서로 독립이다.

$\boxed{2}$ $P(A)=0$ 또는 $P(B)=0$이면, A와 B는 서로 독립이다.

$\boxed{3}$ $P(A)>0$, $P(B)>0$이고 A와 B가 서로 배반이면, A와 B는 서로 종속이다.

Example

어떤 공장의 생산라인 공정은 서로 독립적인 두 부분의 기계장치 A, B로 구성되어 있고, 기계장치 A, B가 고장 날 확률은 각각 20%, 15%일 때, 두 기계장치 가운데 적어도 한 기계장치가 고장 날 확률을 구하여라.

Example

어떤 프로그래머는 컴퓨터를 이용하여 작업한 파일을 하드와 USB에 백업을 한다. 하드에 백업 할 때 파일이 훼손될 확률이 1.2%, USB에 백업할 때 파일이 훼손될 확률이 2.5%라 하자. 두 방법으로 백업하는 사건이 서로 독립일 때, 이 프로그래머가 적어도 하나의 훼손되지 않은 파일을 가질 확률을 구하여라.

> ***Definition***
>
> 서로 다른 n개의 사건 $A_1, A_2, \cdots, A_n$에 대하여
>
> (1) 모든 $i\neq j\ (1\le i,j\le n)$에 대하여
>
> $$P(A_i\cap A_j)=P(A_i)P(A_j)$$
>
> 이면, 사건 $A_1, A_2, \cdots, A_n$을 **쌍마다 독립**(pairwisely independent)이라 한다.
>
> (2) 모든 $J=\{j_1, j_2, \cdots, j_k\}\subset\{1, 2, \cdots, n\}$에 대하여,
>
> $$P\left(\bigcap_{j=j_1}^{j_k} A_j\right)=\prod_{j=j_1}^{j_k} P(A_j)$$
>
> 이면, 사건 $A_1, A_2, \cdots, A_n$을 **독립**(independent)이라 한다.

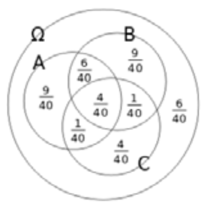
Example

동전을 세 번 던지는 게임에서 첫 번째 던진 동전이 앞면일 사건을 A_1, 두 번째 던진 동전이 앞면일 사건을 A_2, 두 번만 연속해서 앞면이 나올 사건을 A_3라 할 때, 사건 A_1, A_2, A_3가 독립인지 확인하여라.

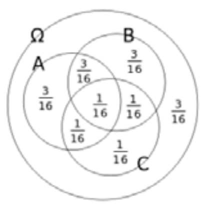
Example

다음 표본공간 Ω의 사건 A, B, C가 독립인지 확인하여라.

(1)



(2)



Theorem **전확률 공식** (formula of total probability)

사건 $A_1, A_2, \cdots, A_n$이 표본공간 Ω의 분할이고

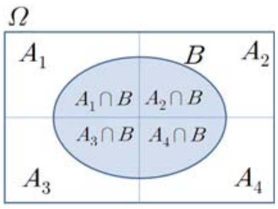
$$P(A_i) > 0 \ (i = 1, 2, \cdots, n)$$

일 때, 임의의 사건 B에 대하여

$$P(B) = \sum_{i=1}^{n} P(A_i) P(B \mid A_i)$$

이다.

(proof)



Theorem **베이즈 정리** (Bayes' Theorem)

사건 $A_1, A_2, \cdots, A_n$이 표본공간 Ω의 분할이고

$$P(A_i) > 0 \ (i = 1, 2, \cdots, n)$$

일 때, $P(B) > 0$인 사건 B에 대한 사건 A_i의 조건부확률은 다음과 같다.

$$P(A_i \mid B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i) P(B \mid A_i)}{\sum_{j=1}^{n} P(A_j) P(B \mid A_j)}$$

참고 $P(B)$: **전확률**

결과로서 B가 관측될 확률

$P(A_i)$: **사전확률** (prior probability)

결과 B가 아직 관측되지 않은 단계에서 원인 A_i의 확률

$P(A_i \mid B)$: **사후확률** (posterior probability)

결과로서 B가 관측되었을 때, 원인이 A_i일 확률

$P(B \mid A_i)$: **가능도** (likelihood)

원인이 A_i일 때 결과로서 B를 관측할 확신의 정도

Example

어떤 공장의 생산품은 세 생산라인 A_1, A_2, A_3을 통과해서 만든다고 한다. A_1에서는 생산품의 50%, A_2에서는 25%, A_3에서는 25%를 맡고 있고, A_1에서 불량률이 2%, A_2에서 불량률이 2%, A_3에서 불량률이 4%라 한다.

(1) 생산품 중 하나를 임의로 선정했을 때, 이 제품이 불량품일 확률을 구하여라.

(2) 임의로 선정된 제품이 불량품이였을 때, A_1에서 만들어졌을 확률을 구하여라.

Example

의학 보고서에 따르면 전체 국민의 7%가 폐질환을 앓고 있으며, 그들 중 85%가 흡연가라고 한다. 그리고 폐질환을 갖지 않은 사람 중에 25%가 흡연가라 한다.

(1) 임의로 선정한 사람이 흡연가일 확률을 구하여라.

(2) 임의로 선정한 사람이 흡연가라 할 때, 이 사람이 폐질환을 앓고 있을 확률을 구하여라.