Introduction to Statistic

Written by Eun Taek Kang*[1]

[1] *Department of Physics, Sogang University, Seoul 04107, Korea*

Summer 2025, Sogang University

Abstract

본 문서는 서강대학교의 2025년 하계학기 통계학입문 강의의 내용을 필기하고 정리한 노트입니다. 모든 내용은 임경필 교수님의 강의를 바탕으로 작성되었습니다.

*email: etkang03@gmail.com

1 확률

1.1 확률의 정의

확률현상, 확률실험

- 확률현상(random phenomenon) : 불확실성에 의해서 좌우되는 현상
- 확률실험(random experiment) : 결과가 불확실성에 의해 좌우되는 실험이나 조사

Example 1.1 (확률현상의 예)**.** 확률 현상의 예시들은 다음과 같다.

1. 동전이나 주사위를 던질 때 결과가 나타나는 현상
2. 각 지역에서의 연 강수량의 변동 현상
3. 수시로 달라지는 주가의 변동 현상
4. 월 생산품에 포함되는 불량품(수)의 출현 현상

표본공간, 표본점, 사건

1. 표본공간(sample space; Ω) : 한 확률실험에서 얻어질 수 있는 가능한 모든 결과의 집합. 표본공간 = 전공간
2. 표본점(sample point; ω_i) : 표본공간의 원소 하나
3. 사건(event) = 사상 : 표본공간의 부분집합
4. 근원사건(elementary event)= 단순사건(simple event) : 표본점 하나하나로 이루어지는 사건

Example 1.2. 확률실험에 대한 표본공간의 예시는 다음과 같다.

1. 하나의 동전을 던지고 윗면에 나타나는 상태를 관찰하는 확률실험: (단, Heads(앞면-그림), Tails(뒷면-금액))

   표본공간 : $\Omega = \{H, T\}$
2. 하나의 주사위를 던지고 윗면에 나타나는 점의 개수를 관찰하는 확률실험:

   표본공간 : $\Omega = \{1, 2, 3, 4, 5, 6\}$

우리는 셀 수 있는 공간(countable)뿐 아니라 셀 수 없는 공간(uncountable)에서의 확률도 고려해야 한다. 이때 등장하는 것이 **기하적 확률**(Ex. 길이, 넓이, 부피)이다.[1]

주사위를 굴려서 짝수가 나온다는 확률을 잴 때, {2}, {4}, {6}으로 쪼개지 {2, 4}, {4, 6}으로 굳이 쪼개어 생각하지 않는다. 따라서 이후 나올 배반사건으로 쪼개서 생각하는 것이 상당히 중요한 사고로 작용한다.

[1] 이런 부분은 실해석학(Real Analysis)의 공부에 도움이 된다.

사건의 결합(combination of events)

1. 여사건: E_1^c
2. 합사건: $E_1 \cup E_2$
3. 교사건: $E_1 \cap E_2 = E_1E_2$[2]
4. 차사건: $E_1 - E_2$
5. 무한가산합사건: $\bigcup_{i=1}^{\infty} E_i = \{\omega \in \Omega \mid \exists\, i \in \mathbb{N},\ \omega \in E_i\}$
6. 무한가산교사건: $\bigcap_{i=1}^{\infty} E_i = \{\omega \in \Omega \mid \forall\, i \in \mathbb{N},\ \omega \in E_i\}$

사전확률, 사후확률

- 사전확률 : 확률실험에서 모든 결과가 같은 가능성을 갖고 일어난다고 하자. 이 실험에서 일어날 수 있는 가능한 모든 결과의 수가 N, 또 한 사건 E를 일어나게 하는 결과의 수가 n이면, 이때 사건 E가 일어날 확률은 $P(E) = n/N$이다.
- 사후확률 : 한 확률실험을 반복한 총 횟수를 N 이라 하고, 또 그 중 사건 E를 일어나게 한 실험의 횟수를 n 이라 하면, 이때 사건 E가 일어날 확률은 $P(E) = \lim_{N\to\infty} n/N$이다.

배반사건(mutually exclusive events)

Definition 1.3. 두 사건 E_1과 E_2가 $E_1 \cap E_2 = \emptyset$을 만족할 때, E_1과 E_2는 상호 배반사건이라 한다.

경우의 수와 확률을 분석할 때, 상호 배반사건 관계에 있는 것들로 쪼개주는 것이 중요하다. 윷놀이 예시가 있는데 한 번쯤 분석해보도록 합시다.

공리론적 확률
정의역때문에 설명이 어려워서 그냥 대충 받아들이고 가도 무방하나, 실해석학할 때 도움된다. (실사멸하;;)

Definition 1.4. Ω를 확률실험의 표본공간이라 하고, $\mathcal{F}$를 Ω 위에서 정의된 사건들의 'σ-집합체'라고 할 때, 집합함수 $P : \mathcal{F} \to \mathbb{R}$ ($\mathcal{F} \subset P(\Omega)$)가 다음 세 가지 조건을 만족할 때 P를 확률함수(probability function)라 하고, $(\Omega, \mathcal{F}, P)$를 확률공간 (probability space)이라 한다.

(A1) $\forall E \in \mathcal{F},\ P(E) \geq 0$

(A2) 임의의 배반사건인 $\{E_n\}_{n=1}^{\infty} \subset \mathcal{F}$에 대하여,

$$P\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} P(E_n)$$

(A3) $P(\Omega) = 1$

상호 배반 합집합의 경우, (A2)의 $\cup$의 중간에 점을 찍어 표현하기도 한다. 이러한 공리론적 확률덕분에 배반사건에 대해서 합을 구하는 것은 상당히 쉽게 된다.

[2]중간의 교집합 기호를 생략할 수 있다. Cartesian product가 아님에 유의하자.

σ-집합체(σ-field), σ-집합대수(σ-algebra)

Definition 1.5. Ω를 확률실험의 표본공간이라 하고, $\mathcal{F}$를 Ω의 부분 집합족이라 할 때, 다음 세가지 조건을 만족할 때 $\mathcal{F}$를 Ω 위에서의 σ-집합체 또는 σ-집합대수라 한다.

(i) $E \in \mathcal{F} \Rightarrow E^c \in \mathcal{F}$

(ii) 임의의 $\mathcal{F}$의 가산무한 사건열 $\{E_n\}_{n=1}^{\infty}$에 대하여

$$\bigcup_{n=1}^{\infty} E_n \in \mathcal{F}$$

(iii) $\emptyset \in \mathcal{F}$

Theorem 1.6 (σ-집합체의 기본 성질)**.** $\mathcal{F}$를 Ω 위에서의 σ-집합체라 하자.

1. $\Omega \in \mathcal{F}$
2. 임의의 $\mathcal{F}$의 가산무한 사건열 $\{E_n\}_{n=1}^{\infty}$에 대하여

$$\bigcap_{n=1}^{\infty} E_n \in \mathcal{F}$$

Definition 1.7 (Borel σ-집합체)**.** $\mathbb{R}$의 부분집합들의 집합족 $\{(-\infty, x] \mid x \in \mathbb{R}\}$를 포함하는 $\mathbb{R}$ 위에서의 가장 작은 σ-집합체 $\mathcal{B}$를 Borel σ-집합체라고 한다. Borel σ-집합체 $\mathcal{B}$는 $\mathbb{R}$의 모든 형태의 구간, 즉, 다음과 같은 형태의 구간을 포함한다.

$$(-\infty, a], \quad (-\infty, a), \quad [a, \infty), \quad (a, \infty), \quad (a, b), \quad (a, b], \quad [a, b), \quad [a, b]$$

여기서,

$$(-\infty, a) = \bigcup_{n \in \mathbb{N}} \left(-\infty, a - \frac{1}{n}\right]$$

Theorem 1.8. $(\Omega, \mathcal{F}, P)$가 확률공간일 때,

1. 임의의 상호 배반인 $\mathcal{F}$의 유한 사건열 $\{E_1, E_2, \ldots, E_n\}$에 대하여

$$P(E_1 \cup E_2 \cup \cdots \cup E_n) = P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i)$$

2. $\forall E \in \mathcal{F}, \quad P(E^c) = 1 - P(E)$
3. $\forall E \in \mathcal{F}, \quad 0 \leq P(E) \leq 1$
4. $P(\emptyset) = 0$

Proof. $(\Omega, \mathcal{P}, P)$: 확률 공간
(1) : 증명은 생략한다.
(2) : $E \cup E^c = \Omega$이고, $E \cap E^c = \emptyset$ (상호 배반)이므로, (A2)에 의해 $P(E \cup E^c) = P(E) + P(E^c)$이고, (A3)에 의해 $P(\Omega) = 1$이다. $P(E \cup E^c) = P(\Omega)$이므로, $P(E^c) = 1 - P(E)$이 증명된다.
(3) : (A1)에 의하여 $P \geq 0$이 성립한다. (2)에서 얻었던 점에 의하여 $P(E) = P(\Omega) - P(E^c)$였다. 따라서, 부등식을 세울 수 있다. $P(E) \leq P(\Omega) = 1$ 따라서, 두 부등식을 종합하면 $0 \leq P(E) \leq 1$이 증명된다.
(4) : $\Omega = \Omega \cup \emptyset$이다. 둘은 상호 배반이므로 $P(\Omega) = P(\Omega) + P(\emptyset)$이다. 따라서 $P(\emptyset) = 0$이 증명되었다. □

[Lecture 1 to here]