Mathematical Statistics

- Implementation with R -

Chap 1. Probability

Some Definitions

Sample space(Ω or S) [표본공간]

- The collection of all outcomes from an experiment
- 어떤 실험 또는 시행에서 나올 수 있는 모든 가능한 결과의 집합

예1) A rat is selected at random from a cage, and its sex is determined.
The set of possible outcomes is female and male. Thus,

$$\Omega = \{female, male\} = \{F, M\}$$

예2) A fair coin is flipped until heads is observed.
Let k denote the number of trials required. Then,

$$\Omega = \{k : k = 1, 2, , \cdots\}$$

예3) Results of tossing a fair dice twice.

$$\Omega = \{(1,1),(1,2),\cdots,(6,5),(6,6)\}$$

예4) Results of flipping two fair coins.
(동전 앞면 : H(head), 뒷면 : T(tail))

$$\Omega = \{(H,H),(H,T),(T,H),(T,T)\}$$

Event [사건]

- The part of sample space S
- 표본공간의 원소 또는 부분 집합으로 관심의 대상이 되는 결과의 집합

1.1 Properties of Probability

Probability [확률]

- Set function that assigns to each event a number
- Express the level of trust that an event occur
- Relative frequency associated with an event
- Numerical expression of the possibility of an event

Elements of Probability

- Sample space(S)
- Event(A)
- Probability(P)

$\Rightarrow$ (Ω, A, P)

1.1 Properties of Probability

Axiom of Probability [AP: 확률의 공리]

(i) For sample space Ω, $P(\Omega) = 1$

(ii) For each event A, $0 \leq P(A) \leq 1$

(iii) For disjoint event $A_1, A_2, \cdots$, (that is, for $i \neq j$, $A_i \cap A_j = \phi$)

$$P\left(\cup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

i.e. for finite case, $P\left(\cup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$

1.1 Properties of Probability

Theorem 1.1

(a) Let A^c be the compliment of A. Then, $P(A^c) = 1 - P(A)$

pf) Since $A \cup A^c = \Omega$, $P(\Omega) = P(A) + P(A^c) = 1$ [by AP (i) & (ii)]

(b) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. (Additivity)

pf) $A \cup B = A \cup (A^c \cap B)$, $A \cap (A^c \cap B) = \phi$

$\Rightarrow P(A \cup B) = P(A) + P(A^c \cap B)$

$B = (A \cap B) \cup (A^c \cap B)$, $(A \cap B) \cap (A^c \cap B) = \phi$

$\Rightarrow P(B) = P(A \cap B) + P(A^c \cap B)$ [by AP (iii)]

1.1 Properties of Probability

(c) When $A \subset B$, $P(A) \le P(B)$

pf) $A \subset B$ implies $B = A \cup (A^c \cap B)$.

Then, $P(B) = P(A) + P(A^c \cap B)$ & $P(A^c \cap B) \ge 0$ [by AP (ii) & (iii)]

Exercise

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

proofs) H.W

1.2 Methods of Enumeration

Additivity principle(합의 원리)

- The number of cases choosing one possible outcome among m outcomes is m.
- m개의 가능한 방법 중에서 하나를 선택하는 경우의 수는 m개

Multiplication principle(곱의 원리)

- Experiment $M1$ has n_1 outcomes, $M2$ has n_2 possible outcomes. Then, the composite experiment $M1M2$ has $n_1 n_2$ possible outcomes.

1.2 Methods of Enumeration

Combination(조합)

- 서로 다른 n개의 다른 대상 중에서 r개를 뽑는 경우(${}_nC_r$ or $\binom{n}{r}$)

$${}_nC_r = \binom{n}{r} = \frac{n!}{r!(n-r)!}$$

$x! = x(x-1)(x-2)\cdots 2\cdot 1, 0! = 1$로 정의

1.2 Methods of Enumeration

Permutation(순열)

- 서로 다른 n개 중에 중복 없이 r개를 뽑아서 배열하는 경우($_nC_r$)
- n개의 대상 중에 r개를 뽑고($_nP_r$), 뽑힌 각 경우에 대해 다시 이 r개를 배열하는 경우의 수($r!$)를 곱한 것

$$_nP_r = {}_nC_r \times \mathrm{r}! = \frac{n!}{r!(n-r)!} \times r! = n(n-1)(n-2)\cdots(n-r+1)$$

1.2 Methods of Enumeration

[예 1.1]

(a) 10명의 학생 중에서, 2명의 학생 대표를 뽑는 경우의 수

(b) 10명의 학생 중에서 회장 1명, 부회장 1명을 뽑는 방법의 수

[예 1.2]

(a) “MOUSE” 배열하는 경우의 수

(b) “EXAMPLE” 배열하는 경우의 수

(c) “STATISTICS” 배열하는 경우의 수

1.2 Methods of Enumeration

[예 1.3] 경로관련 예제

가로줄 각 칸 : a, a, a, a, a, 세로줄 각 칸 : b, b, b, b

(a) 전체 경우의 수

(b) C를 거치는 경우의 수

[예 1.4] 남학생 : 8명, 여학생 : 6명.
4명을 선택할 때 남학생이 3명일 확률

1.2 Methods of Enumeration

참고

이항전개

- $(a+b)^n = \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k}$, $a^k b^{n-k}$의 이항계수(Binomial Coefficient)
- $\binom{n}{k}$ & $\sum_{k=0}^{n}\binom{n}{k} = (1+1)^n = 2^n$

다항전개

- $(a_1 + a_2 + \cdots + a_k)^n = \sum_{r_1+\cdots+r_k=n}\binom{n}{r_1 r_2 \cdots r_k} a_1^{r_1} a_2^{r_2} \cdots a_k^{r_k}$,
  $a_1^{r_1} a_2^{r_2} \cdots a_k^{r_k}$의 다항계수(Multinomial Coefficient)
- $\binom{n}{r_1 r_2 \cdots r_k}$, 여기서 $\binom{n}{r_1 r_2 \cdots r_k} = \frac{n!}{r_1! r_2! \cdots r_k!}$

중복순열

- 서로 다른 n개 중에서 중복을 허용하여 r개 뽑아 배열하는 경우의 수

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

예1) 다섯 통의 편지 1, 2, 3, 4, 5를 세 개의 우체통 A, B, C에 넣는 방법의 수

예2) 여섯 개의 수 1, 2, 3, 4, 5, 6 중에서 중복을 허용하여 네 개의 수를 택하여 만든 네 자리 정수 중에서 1234 보다 큰 것의 개수?

1.2 Methods of Enumeration

예3) 6층 건물의 엘리베이터에 4명이 타고 1층을 출발했는데 이들이 각 층에서 내릴 수 있는 모든 경우의 수?

예4) 다섯 개의 수 1, 2, 3, 4, 5 중에서 세 개의 수를 택하여 만든 세 자리 정수의 개수? 단, 1과 2는 중복하여 쓸 수 있고 3, 4, 5는 중복하여 쓸 수 없다.

중복조합

- 서로 다른 n개의 원소에서 중복을 허락하여 r개를 뽑는 경우의 수

: ${}_nH_r = {}_{n+r-1}C_r$

Exercise or Homework

1. How many different signals can be made using four flags of different colors on a vertical flagpole if exactly three flags are used for each signal?

2. How many four-letter code words are possible using the letters in HOPE if
   (a) The letters may not be repeated
   (b) The letters may be repeated

1.2 Methods of Enumeration

3. How many solutions with non-negative integers for the equation $x + y + z = 10$?

4. How many solutions with positive integers for the equation $x + y + z = 10$?

5. How many solutions with non-negative integers for the equation $x + y + 4z = 10$?

1.3 Conditional Probability

Conditional probability

- Probability of an event under the condition that a specific event occurs.

-예 : 학생 30명의 성별과 혈액형 자료

Def 1.1

The conditional probability of an event B, given that event A has occurred, is defined by :

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

provided that $P(A) > 0$

1.3 Conditional Probability

Meaning

$n(\Omega)$: # of elements in Ω

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

[예 1.5] (동전 3번 던지기)

사건 A = 앞면이 2회 이상 나오는 사건
사건 B = 앞면이 1회 이상 나오는 사건

참고

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

Theorem 1.2 [Bayes Theorem]

$B_1, B_2, \cdots, B_n$: disjoint, $\cup_{i=1}^{n} B_i = \Omega$, $P(B_i) > 0$, $i = 1, \cdots, n$

$$P(B_i|A) = \frac{P(B_i \cap A)}{P(A)} = \frac{P(B_i) \times P(A|B_i)}{P(B_1) \times P(A|B_1) + P(B_2) \times P(A|B_2) + \cdots + P(B_n) \times P(A|B_n)}$$

참고

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + \cdots + P(A \cap B_n)$$

$$= P(B_1) \times P(A|B_1) + \cdots + P(B_1) \times P(A|B_1)$$

1.3 Conditional Probability

[예 1.6] 첫 번째 항아리 - 흰색 공 : 3개, 검정색 공 : 2개
두 번째 항아리 – 흰색 공 : 1개, 검정색 공 : 2개
첫 번째 항아리에서 공 1개를 임의로 선택 하여 두 번째 항아리에 옮긴 후 두 번째 항아리에서 임의로 공 한 개 선택.

(a) 두 번째 항아리에서 꺼낸 공이 흰색일 확률

(b) 두 번째 항아리에서 꺼낸 공이 흰색일 때 첫 번째 항아리에서 꺼낸 공이 흰색이었을 확률

1.3 Conditional Probability

[예 1.7] 컴퓨터 C에서 컴퓨터 D로 ‘0’ 또는 ‘1’의 정보를 전송.

컴퓨터 C에서 ‘0’을 전송했을 때 D에서 ‘1’로 잘못 인식할 확률 : 0.05,
컴퓨터 C에서 ‘1’을 전송했을 때 D에서 ‘0’으로 잘못 인식할 확률 : 0.03.
컴퓨터 C에서 ‘0’을 전송할 확률은 0.6, ‘1’을 전송할 확률은 0.4.

(a) 컴퓨터 C에서 정보를 전송했을 때 D에서 ‘0’으로 인식할 확률?

(b) 컴퓨터 D에서 ‘0’으로 인식했을 때 C에서 전송한 정보가 ‘0’이었을 확률?

1.4 Independency

Def 1.2 Independent event (독립사건)

두 사건 A, B 에 대해, $P(B|A) = P(B)$ 또는 $P(A|B) = P(A)$

$\Leftrightarrow$ A와 B가 서로 독립.

독립이 아닌 경우 $\Rightarrow$ 서로 종속(dependent)

Multiplicative rule (확률의 곱셈법칙)

임의의 두 사건 A, B에 대하여

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

1.4 Independency

독립사건의 곱셈법칙

서로 독립인 두 사건 A, B에 대하여

$$P(A \cap B) = P(A) \times P(B).$$

[예 1.8] 52장의 카드 중 임의로 한 장을 선택.

스페이드일 사건 : A, 에이스일 사건 : B,

스페이드 에이스일 사건 : C

(a) A와 B는 서로 독립인가?

(b) A와 C는 서로 독립인가?

1.4 Independency

[추가예제] 주머니 안에 50개의 제비가 있다. 이 중에서 2개는 당첨금을 받을 수 있는 것이고, 나머지는 무효 제비이다. 이제 두 사람이 제비를 뽑는데 처음 뽑는 사람이 당첨 제비를 뽑는 사건을 A, 두 번째 사람이 당첨 제비를 뽑는 사건을 B라 하면 두 사건 A와 B는 서로 독립인가?

풀이] 두 가지 상황에 따라 답이 달라짐

① 뽑은 제비를 주머니에 되돌려 넣지 않는 경우

$$P(B) = \frac{2}{50} \times \frac{1}{49} + \frac{48}{50} \times \frac{2}{49} = \frac{98}{2450} = \frac{1}{25}$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{2}{50} \times \frac{1}{49}}{\frac{2}{50}} = \frac{1}{49}$$

$P(B) \neq P(B|A)$이므로 두 사건 A와 B는 서로 독립이 아님

1.4 Independency

② 뽑은 제비를 주머니에 되돌려 넣는 경우

$$P(B) = \frac{2}{50} \times \frac{2}{50} + \frac{48}{50} \times \frac{2}{50} = \frac{100}{2500} = \frac{1}{25}$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{2}{50} \times \frac{2}{50}}{\frac{2}{50}} = \frac{1}{25}$$

$P(B) = P(B|A)$이므로 두 사건 A와 B는 서로 독립

1.4 Independency

Def 1.3 Independence of several events

$A_1, A_2, \cdots, A_n$: 서로 독립

$\Leftrightarrow$ $\{1, 2, \cdots, n\}$의 임의의 부분집합 $\{i_1, i_2, \cdots, i_k\}$에 대해

$$P\left(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}\right) = P(A_{i_1}) \times P(A_{i_2}) \times \cdots \times P(A_{i_k})$$

1.4 Independency

Example : Independence of three events

사건 A, B, C 독립

$$\Leftrightarrow \begin{cases} (i)\ \ P(A \cap B) = P(A) \times P(B) \\ \qquad P(A \cap C) = P(A) \times P(C) \\ \qquad P(B \cap C) = P(B) \times P(C) \\ (i\ )\ P(A \cap B \cap C) = P(A) \times P(B) \times P(C) \end{cases}$$

Note