기계학습

베이즈 정리

함건희

2021년 2학기

한국외국어대학교

들어가며

확률이란?

- Probability is a measure quantifying the likelihood that events will occur. (wikipedia)
- 확률은 어떤 <u>사건</u>의 발생 <u>가능성</u>을 숫자로 표현한 값이다.

확률실험(Random experiment)과 표본공간(Sample space)

확률실험과 표본공간

- 시행하기 전에는 확실히 예측할 수 없는 결과를 유발하는 행위 또는 과정을 **확률실험**이라고 한다.
- 어떤 실험에서 발생 가능한 모든 경우들의 집합을 **표본공간**이라고 한다.

Example 1. **동전을 두 번 던지는 실험**

- 표본공간 $\mathrm{S} = \{(\text{앞면}, \text{앞면}), (\text{앞면}, \text{뒷면}), (\text{뒷면}, \text{앞면}), (\text{뒷면}, \text{뒷면})\}$

Example 2. **주사위를 한 번 던지는 실험**

- 표본공간 $\mathrm{S} = \{1, 2, 3, 4, 5, 6\}$

사건(Event)과 여사건(Complementary event)

사건

- 표본공간에 속한 결과들 중에서 어떤 성질을 만족하는(관심있는) 결과들의 모임을 **사건**이라고 한다. ※ 표본공간의 부분집합
- 표본공간의 원소 한개로 구성된 사건을 단일사건, 두개 이상이 원소로 구성된 사건을 복합사건이라고 한다.

여사건

- 관심있는 사건을 A라고 할 때, 사건 A에 포함되지 않는 표본공간의 원소들로 이루어진 집합을 **여사건**이라고 하며 A^c로 표시한다.

사건(Event)과 여사건(Complementary event)

Example 1. **동전을 두 번 던지는 실험** (revisited.)

- 표본공간 $\mathrm{S} = \{(\text{앞면}, \text{앞면}), (\text{앞면}, \text{뒷면}), (\text{뒷면}, \text{앞면}), (\text{뒷면}, \text{뒷면})\}$
- 첫 시행의 결과가 앞면이 나오는 사건을 A라 할 때,
  $\mathrm{A} = \{(\text{앞면}, \text{앞면}), (\text{앞면}, \text{뒷면})\}$로 나타낼 수 있다.
- 이때, A의 여사건은 $\mathrm{A}^{\mathrm{c}} = \{(\text{뒷면}, \text{앞면}), (\text{뒷면}, \text{뒷면})\}$이다.

Example 2. **주사위를 한 번 던지는 실험** (revisited.)

- 표본공간 $\mathrm{S} = \{1, 2, 3, 4, 5, 6\}$
- 주사위의 눈이 홀수가 나오는 사건을 B라 할 때, $\mathrm{B} = \{1, 3, 5\}$로 나타낼 수 있다.
- 이때, B의 여사건은 $\mathrm{B}^{\mathrm{c}} = \{2, 4, 6\}$이다.

수학적 개념의 확률

- 임의의 사건 A에 대하여

$$\mathrm{P(A)} = \frac{\text{사건 } A\text{에 속하는 단일사건 수}}{\text{표본공간 전체 단일사건수}}$$

Example 2. **주사위를 한 번 던지는 실험** (revisited.)

- 표본공간 $\mathrm{S} = \{1, 2, 3, 4, 5, 6\}$
- B: 홀수 눈이 발생하는 사건, $\mathrm{B} = \{1, 3, 5\}$
- 사건 B의 확률 $\mathrm{P(B)}$와 여사건의 확률 $\mathrm{P(B^c)}$는 다음과 같이 구할 수 있다.

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

$$\mathrm{P(B^c)} = 1 - \mathrm{P(B)} = \frac{1}{2}$$

확률의 공리(Axioms of probability)

확률의 공리

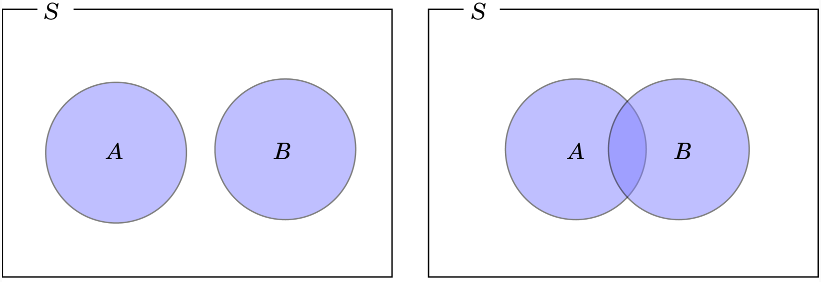
- 임의의 사건 A에 대하여

1. $0 \leq P(A) \leq 1$
2. $P(S) = 1$
3. 서로 **배반**인 사건 $A_i,\ i = 1, 2, \ldots$에 대하여

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

배반사건(Exclusive events)

- 동시에 발생할 수 없는 사건을 **배반사건**이라 한다.
- 만약 두 사건 A와 B가 배반사건이면 서로 공통된 표본공간의 원소를 갖지 않으며 두 사건의 교집합은 공집합이다.$(\mathrm{A} \cap \mathrm{B} = \emptyset,\ \mathrm{P}(\mathrm{A} \cap \mathrm{B}) = 0)$



[그림1] 서로 배반인 사건(좌)과 서로 배반이 아닌 사건(우)

배반사건(Exclusive events)

Example 2. **주사위를 한 번 던지는 실험** (revisited.)

- A: 4 이상의 눈이 발생하는 사건, $A = \{4, 5, 6\}$
- B: 홀수 눈이 발생하는 사건, $B = \{1, 3, 5\}$
- C: 짝수 눈이 발생하는 사건, $C = \{2, 4, 6\}$
- $B \cap C = \emptyset$이므로 사건 B와 C는 서로 배반사건이다.

조건부 확률(Conditional probability)

조건부 확률

- 사건 A가 이미 일어났다는 가정을 전제로 사건 B가 발생할 확률을 **조건부 확률** 이라고 하며 $\mathrm{P(B|A)}$로 표기한다.
- $\mathrm{P(A)} > 0$일때, 조건부 확률은 다음과 같이 나타낼 수 있다.

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

- 만약 사건 A와 B가 서로 배반사건이라면 $\mathrm{P(B|A)} = 0$이다.

조건부 확률(Conditional probability)

Example 2. **주사위를 한 번 던지는 실험** (revisited.)

- 표본공간 $\mathrm{S} = \{1, 2, 3, 4, 5, 6\}$
- A: 4 이상의 눈이 발생하는 사건, $\mathrm{A} = \{4, 5, 6\}$
- B: 홀수 눈이 발생하는 사건, $\mathrm{B} = \{1, 3, 5\}$
- $\mathrm{A} \cap \mathrm{B} = \{5\}$
- 주사위의 눈이 4이상의 숫자가 나왔을 때, 그 눈이 홀수일 확률은 다음과 같다.

$$\mathrm{P}(\mathrm{B}|\mathrm{A}) = \frac{\mathrm{P}(\mathrm{A} \cap \mathrm{B})}{\mathrm{P}(\mathrm{A})} = \frac{1/6}{3/6} = \frac{1}{3}$$

독립(Independent)과 종속(Dependent)

종속

- 두 사건 사이에 밀접한 관계가 있어 한 사건의 발생이 다른 사건의 발생 확률에 영향을 미치게 되면 두 사건은 서로 **종속**이라고 한다.

Example 2. **주사위를 한 번 던지는 실험** (revisited.)

- 표본공간 $\mathrm{S} = \{1, 2, 3, 4, 5, 6\}$
- A: 4 이상의 눈이 발생하는 사건, $\mathrm{A} = \{4, 5, 6\}$
- B: 홀수 눈이 발생하는 사건, $\mathrm{B} = \{1, 3, 5\}$
- 아무런 전제 조건이 없을 때, 사건 B가 발생할 확률 $\mathrm{P(B)} = \frac{1}{2}$
- 사건 A이 발생했을 때, 사건 B가 발생할 확률 $\mathrm{P(B|A)} = \frac{1}{3}$
- 사건 A의 발생이 사건 B의 발생 확률에 영향을 미치므로 사건 A와 B는 서로 종속이다.

독립(Independent)과 종속(Dependent)

독립

- 한사건 발생이 다른 사건의 발생 확률에 영향을 주지 않는다면 두 사건은 서로 <u>독립</u>이라고 한다.

Example 2. **주사위를 한 번 던지는 실험** (revisited.)

- 표본공간 $\mathrm{S} = \{1, 2, 3, 4, 5, 6\}$
- B: 홀수 눈이 발생하는 사건, $\mathrm{B} = \{1, 3, 5\}$
- D: 1 또는 2의 눈이 발생하는 사건, $\mathrm{D} = \{1, 2\}$
- $\mathrm{B} \cap \mathrm{D} = \{1\}$
- 아무런 전제 조건이 없을 때, 사건 D가 발생할 확률 $\mathrm{P(D)} = \frac{1}{3}$
- 사건 B가 발생했을 때, 사건 D가 발생할 확률 $\mathrm{P(D|B)} = \frac{\mathrm{P(B \cap D)}}{\mathrm{P(B)}} = \frac{1/6}{1/2} = \frac{1}{3}$
- 사건 B의 발생이 사건 D의 발생 확률에 영향을 미치지 않으므로 사건 B와 D는 서로 독립이다.

독립(Independent)과 종속(Dependent)

- $\mathrm{P(A)}, P(B) > 0$인 두 사건 A와 B가 독립이기 위한 **필요충분조건**은 다음과 같다.

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

cf. 배반사건: $\mathrm{P(A \cap B) = 0}$

- $\mathrm{P(A), P(B), P(C) > 0}$인 세 사건 $\mathrm{A, B, C}$가 서로 독립이기 위한 **필요충분조건**은 다음과 같다.

1. $\mathrm{P(A \cap B) = P(A)P(B), P(A \cap C) = P(A)P(C), P(B \cap C) = P(B)P(C)}$
2. $\mathrm{P(A \cap B \cap C) = P(A)P(B)P(C)}$

Example 4. **출석률과 학점**

- 아래의 표는 '확률과통계' 수강생 60명의 출석률과 성적에 대한 결과이다.

[표1] '확률과통계' 수강생 60명의 출석률과 성적

	A학점(A)	B학점(B)	C학점 이하(C)
출석률 90% 이상 (R)	0.20	0.30	0.10
출석률 90% 이하 (R^c)	0.00	0.20	0.20

1. 임의로 한 학생을 선택했을 때, 그 학생의 학점이 A일 확률은?
2. 임의로 한 학생을 선택했을 때, 그 학생의 출석률이 90% 이상일 확률은?
3. 임으로 선택한 학생의 출석률이 90% 이상일 때, 그 학생의 학점이 A일 확률은?
4. 출석률 90% 이상인 사건과 A학점을 받는 사건은 서로 독립인가?

독립(Independent)과 종속(Dependent)

[solution]

1. $\mathrm{P(A)} = 0.20 + 0.00 = 0.20$
2. $\mathrm{P(R)} = 0.20 + 0.30 + 0.10 = 0.60$
3. $\mathrm{P(A|R)} = \frac{\mathrm{P(A \cap R)}}{\mathrm{P(R)}} = \frac{0.20}{0.60} \approx 0.33$
4. $\mathrm{P(A)} = 0.2 \neq \mathrm{P(A|R)} \approx 0.33$이므로 사건 A와 R은 서로 종속이다.

확률의 법칙

합의 법칙(Addition rule)

- 임의의 두 사건 A와 B에 대하여
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
- 두 사건 A와 B가 배반이라면 $P(A \cap B) = 0$이므로
$$P(A \cup B) = P(A) + P(B)$$

Example 5. **주사위를 두 번 던지는 실험**

- 표본공간 $S = \{(1,1), (1,2), \ldots, (6,5), (6,6)\}$

  ※ 총 36(6×6)가지 단일사건이 존재한다.
- A: 두 눈의 합이 8인 사건 = $\{(2,6), (3,5), (4,4), (5,3), (6,2)\}$
- B: 두 눈의 곱이 12인 사건 = $\{(2,6), (3,4), (4,3), (6,2)\}$
- $A \cap B = \{(2,6), (6,2)\}$
- $P(A) = \frac{5}{36}$, $P(B) = \frac{4}{36}$, $P(A \cap B) = \frac{2}{36}$
$$P(A \cup B) = \frac{5}{36} + \frac{4}{36} - \frac{2}{36} = \frac{7}{36}$$

확률의 법칙

곱의 법칙(Multiplication rule)

- 임의의 두 사건 A와 B에 대하여

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

- 두 사건 A와 B가 독립이라면 $P(A|B) = P(A)$, $P(B|A) = P(B)$ 이므로

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

Example 5. 주사위를 두 번 던지는 실험 (revisited.)

- 표본공간 $S = \{(1,1), (1,2), \ldots, (6,5), (6,6)\}$
- A: 두 눈의 합이 8인 사건 = $\{(2,6), (3,5), (4,4), (5,3), (6,2)\}$
- C: 두 눈이 같은 사건 = $\{(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)\}$
- $P(A) = \frac{5}{36}$, $P(C|A) = \frac{1}{5}$

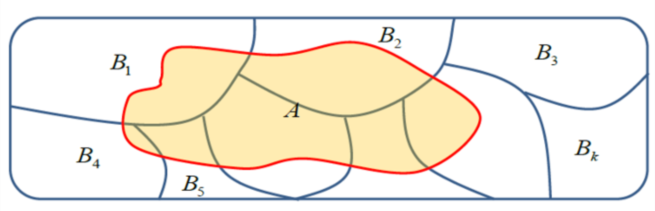
$$P(A \cap C) = P(C|A)P(A) = \frac{1}{5} \times \frac{5}{36} = \frac{1}{36}$$

분할(Partition)

분할

- 사건 $B_1, B_2, \ldots, B_k$ 들이 표본공간 S에 대하여 다음의 조건을 만족한다면, S에 대한 **분할 (partition)**이라 한다.

  1. 모든 $i \neq j$인 i, j에 대해 $B_i \cap B_j = \emptyset$
  2. $\bigcup_{i=1}^{k} B_i = S$



전확률의 정리(Theorem of total probabilities)

전확률의 정리

- 서로 배반인 사건 $B_1, \ldots, B_k$가 표본공간 S의 분할일 때, S에 속하는 임의의 사건 A에 대해 다음이 성립한다.

$$
\begin{aligned}
P(A) &= P(A \cap S) \\
&= P\left(A \cap \left(\bigcup_{i=1}^{k} B_i\right)\right) \\
&= \sum_{i=1}^{k} P\left(A \cap B_i\right) \\
&= \sum_{i=1}^{k} P(A|B_i)P(B_i)
\end{aligned}
$$

전확률의 정리(Theorem of total probabilities)

Example 6. **전확률의 정리 예제**

- 어떤 공장에서 기계 A, B, C 3대가 각각 생산하는 상품이 전체 상품의 50%, 20%, 30%의 비율을 차지한다고 한다. 그리고 이들 기계가 생산한 불량품의 비율은 각각 3%, 5%, 4%라고 한다. 이때, 생산품 중에서 임의로 한 개를 선택했을 경우 그 제품이 불량품일 확률을 구하라.

[solution]

- D: 제품이 불량품인 사건
- $P(A) = 0.5$, $P(B) = 0.2$, $P(C) = 0.3$, $P(D|A) = 0.03$, $P(D|B) = 0.05$, $P(D|C) = 0.04$

$$\begin{aligned}
P(D) =& P(D \cap S) \\
=& P\left(D \cap (A \cup B \cup C)\right) \\
=& P(D \cap A) + P(D \cap B) + P(D \cap C) \\
=& P(D|A)P(A) + P(D|B)P(B) + P(D|C)P(C) \\
=& 0.03 \times 0.5 + 0.05 \times 0.2 + 0.04 \times 0.3 \\
=& 0.037
\end{aligned}$$

베이즈 정리

- $P(B_i) > 0$ $(i = 1, \ldots, k)$인 서로 배반인 사건 $B_1, \ldots, B_k$가 표본공간 S의 분할일 때, $P(A) > 0$인 사건 A에 대해 다음이 성립한다.

$$P(B_j|A) = \frac{P(A \cap B_j)}{P(A)} = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{k} P(A|B_i)P(B_i)}$$

Example 7. **베이즈 정리 예제**

- '기계학습' 수업을 듣는 학생은 1학년 16명, 2학년 22명, 3학년 17명, 4학년 5명으로 이루어져있다. 학기가 끝난 후 조사해보니 A학점을 받은 학생은 각 학년별로 3명씩이었다. 임의로 뽑은 한 학생의 학점이 A일 때, 그 학생이 4학년일 확률은?

베이즈 정리(Bayes' theorem)

[solution]

- A: A학점을 받은 사건
- G_1: 1학년인 사건, G_2: 2학년인 사건, G_3: 3학년인 사건, G_4: 4학년인 사건
- $P(G_1) = \frac{16}{60}$, $P(G_2) = \frac{22}{60}$, $P(G_3) = \frac{17}{60}$, $P(G_4) = \frac{5}{60}$
- $P(A|G_1) = \frac{3}{16}$, $P(A|G_2) = \frac{3}{22}$, $P(A|G_3) = \frac{3}{17}$, $P(A|G_4) = \frac{3}{5}$

$$
\begin{aligned}
P(G_4|A) &= \frac{P(A \cap G_4)}{P(A)} \\
&= \frac{P(A \cap G_4)}{P(A \cap G_1) + P(A \cap G_2) + P(A \cap G_3) + P(A \cap G_4)} \\
&= \frac{P(A|G_4)P(G_4)}{P(A|G_1)P(G_1) + P(A|G_2)P(G_2) + P(A|G_3)P(G_3) + P(A|G_4)P(G_4)} \\
&= \frac{\frac{3}{5} \times \frac{5}{60}}{\frac{3}{16} \times \frac{16}{60} + \frac{3}{22} \times \frac{22}{60} + \frac{3}{17} \times \frac{17}{60} + \frac{3}{5} \times \frac{5}{60}} = \frac{1}{4}
\end{aligned}
$$