기초 확률 이론

결정모형 vs 확률모형

- 결정모형(deterministic model)
  - 어떤 현상에 대한 과학적 연구에 있어서 그 현상이 지니고 있는 성격 또는 특징에 대한 정확한 기술이나 예측을 하기 위한 수리적 모형
  - 예를 들면, 어떤 높은 지점에서 물체를 자유 낙하시켰을 때 시점 t에서의 낙하속도 $v = gt$ (중력 g, 시점 t)
    - ▶ 몇 번이고 물체의 자유낙하를 반복 시행하더라도 중력과 시점이 같으면, 언제나 같은 속도가 구해진다.

결정모형 vs 확률모형

- 확률모형(probability model)
  - 반복시행 때마다 나오는 결과가 찬스(chance)에 의존하는 유형
  - 예를 들면, 동전을 반복하여 10회 던졌을 때 앞면이 나오는 횟수는 고정되어 있는 것이 아니라 찬스(chance)에 의해 결정
  - 통계적 정규성(statistical regularity)을 규명하기 위하여 고려되는 모델

확률모형

확률모형을 전개하는 데 있어서 필요한 용어

- 실험
  어떤 현상의 관찰결과를 얻기 위한 과정 (experiment)

- 표본공간 S
  모든 관찰 가능한 결과의 집합 (sample space)

- 사건
  표본공간의 부분집합 (event)

실험 / 표본공간 / 사건의 예제

1. 동전을 3회 던지는 실험을 시행할 때, 동전의 앞면을 H로, 뒷면을 T로 표시

✓ 8개의 원소로 구성된 표본공간 :

$$S = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$$

✓ 2회 앞면이 나오는 사건 :

$$A = \{HHT, HTH, THH\}$$

2. 어떤 기계의 수명시간에 대한 측정

✓ 적절한 표본공간 :

$$S = \{x|0 \leq x \leq \infty\}$$

✓ 20시간 이상 작동할 사건 :

$$A = \{x|x \geq 20\}$$

사건의 교집합 / 합집합

사건 A와 B가 동일 표본공간 S 상에 정의

- 교집합 $A \cap B$

  사건 A와 B에 동시에 속하는 사건 (intersection)

- 합집합 $A \cup B$

  사건 A 또는 B에 속하는 사건 (union)

사건의 교집합 / 합집합의 예제

- 주사위를 1회 던져 나오는 눈의 수를 관찰
  - ✓ 표본공간 :

    $S = \{1,2,3,4,5,6\}$
  - ✓ 사건 A : 짝수의 눈이 나오는 경우

    $A = \{2,4,6\}$
  - ✓ 사건 B : 3의 배수가 나오는 경우

    $B = \{3,6\}$
  - ✓ $A \cap B$:

    $\{6\}$
  - ✓ $A \cup B$:

    $B = \{2,3,4,6\}$

사건의 상호 배반 / 여집합

사건 A와 B가 동일 표본공간 S 상에 정의

- 상호배반

  사건 A와 B가 공통된 부분이 없을 때, 즉 $A \cap B = \emptyset$ (mutually exclusive)

- 여사건 A^c

  사건 A에 포함되지 않는 모든 S의 원소의 집합 (complement)

사건의 상호 배반 / 여집합의 예제

① 동전을 3회 던지는 실험

✓ 사건 A : 최소한 한 번의 앞면이 나오는 사건

$\{HHH, HHT, HTH, THH, HTT, THT, TTH\}$

✓ 사건 B : 모두 뒷면이 나오는 사건

$\{TTT\}$

✓ 사건 A와 B는 공통부분이 없으므로 상호배반

② 주사위를 던지는 실험에서 나오는 눈의 크기

✓ 표본공간 :

$\{1,2,3,4,5,6\}$

✓ 사건 A : 짝수의 눈

$\{2,4,6\}$

✓ 사건 B : 홀수의 눈

$\{1,3,5\}$

✓ 사건 A는 표본공간에서 사건 B를 제외한 부분이기 때문에 사건 B의 여사건이다.

확률의 종류

- 등확률
- 상대 도수(relative frequency)의 극한
- 주관적 확률 (Bayesian 통계)

등확률

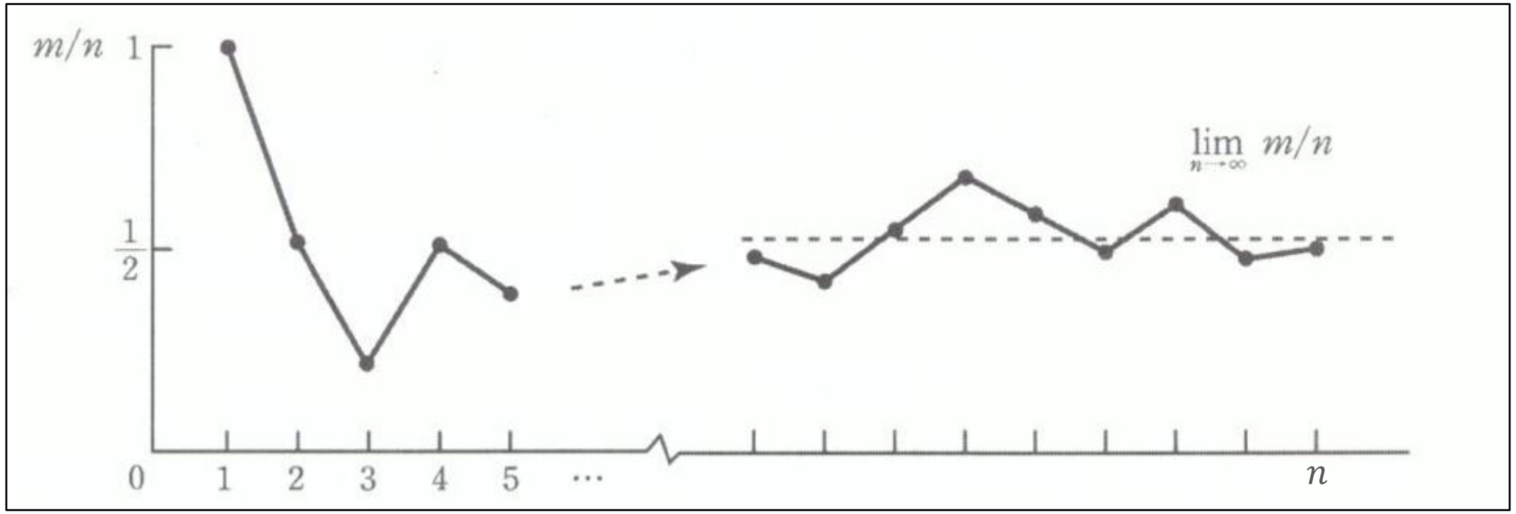
- 확률을 정의하는 가장 오래된 방법
- 표본공간이 유한개(N)의 결과들로 구성되고, 모든 가능한 실험결과들이 등확률로 일어나는 경우에 적용
- 사건 A가 M개의 실험결과를 포함한다면 사건 A가 일어날 확률, $P(A) = \frac{M}{N}$이 된다는 것이다.
- 예를 들어, 동전을 2회 던졌을 때 1개의 앞면이 나올 사건 $A(=\{HT, TH\})$에 대한 확률 $P(A)$
  - 동전을 2회 던졌을 때의 표본공간은 $S = \{HH, HT, TH, TT\}$
  - 동전이 공정한 경우 표본공간에 속한 4개의 원소는 등확률로 일어난다고 가정
  - $P(A) = P(HT, TH) = \frac{2}{4} = \frac{1}{2}$
- 등확률로 확률을 정의하는 것이 곤란한 이유
  - 표곤공간에 속한 원소들이 무한개가 되는 경우가 있다.
  - 모두 같은 확률을 갖는다는 가정이 맞지 않는 경우가 있다.

상대 도수(relative frequency)의 극한

- 가장 널리 받아들여지는 확률의 해석
- 실험을 같은 조건에서 독립적으로 n회 반복하였을 때, 사건 A의 발생횟수를 m이라 하자. $\frac{m}{n}$의 값은 비 결정적이긴 하지만 시행횟수 n이 커짐에 따라 안정될 것으로 예상할 수 있다. 따라서 실험을 무한히 반복한다는 가정하에서 상대도수 $\frac{m}{n}$의 극한값을 사건 A가 일어날 확률 $P(A)$로 정의
- 예를 들어, 어떤 동전을 독립적으로 n회 반복해서 던졌을 때 앞면이 나오는 횟수를 m이라고 하자.
  - $\frac{m}{n}$은 고정된 상수가 아니며, 새로운 실험을 할 때마다 다른 값을 가진다.
  - 만일 n이 커짐에 따라 상대도수 $\frac{m}{n}$이 $\frac{1}{2}$로 수렴하면, 동전을 1회 던졌을 때 앞면이 나올 사건 A의 확률은 $P(A) = \frac{1}{2}$이 된다는 것이다.

상대 도수(relative frequency)의 극한

동전던지기의 독립반복시행에 있어서 앞면이 나오는 사건의 상대도수



- 표본공간이 무한이거나 실험결과들이 등확률로 일어나지 않는 경우에도 적용

주관적 확률 (Bayesian 통계)

- 확률을 불확실성의 측도로 보는 접근방법
- $P(A)$는 A라는 사건이 일어나는 것에 대한 개개인의 믿음의 수준을 표현

확률공리

표본공간 S, S의 부분집합 A에 대하여,

- 임의의 사건 A에 대하여 $P(A) \geq 0$ 이다.
- $P(S) = 1$ 이다.
- 표본공간 S에 정의된 사건열 $A_1, A_2, \ldots$가 있다고 하자. 이제 모든 $i \neq j$ 에 대하여

  $A_i \cap A_j = \emptyset$ 이면 $P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

확률이란 표본공간을 정의역으로 하면서 위의 세공리를 만족하는 함수로 정의

fast campus

확률의 성질

두 개의 사건 A와 B에 대하여 다음과 같은 성질들이 성립한다.

- $P(A^c) = 1 - P(A)$
- $P(\emptyset) = 0$
- $A \subset B$이면 $P(A) \leq P(B)$ 이다.
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

확률의 성질 예제1

- 동전을 3번 던지는 실험에서 동전 앞면이나 뒷면이 나온다.
  - ✓ 표본공간 S의 점들은 등확률(1/8)을 가졌다고 가정
  - ✓ 사건 A : 최소한 1개의 앞면이 나오는 사건

    $P(A) = 1 - P(A^c) = ?$
  - ✓ 사건 B : 꼭 2개의 앞면이 나오는 사건
  - ✓ 사건 C : 첫 번째 동전이 앞면이 나오는 사건

    $P(B \cup C) = ?$

    $P(B) + P(C) - P(B \cap C) = ?$

확률의 성질 예제2

- 주사위를 한 번 던지는 실험에서 1~6사이의 눈의 크기가 나온다.
  - ✓ 표본공간 S의 점들은 등확률(1/6)을 가졌다고 가정
  - ✓ 사건 A : 짝수눈이 나오는 사건

    $P(A) = ?$
  - ✓ 사건 B : 홀수눈이 나오는 사건

    $P(B) = ?$

    $P(A \cap B) = ?$

    $P(A \cup B) = ?$

조건부 확률

사건 A와 B가 표본공간 S상에 정의되어 있으며 $P(B) > 0$ 이라고 하자. 이때 B가 일어났다는 가정하의 사건 A가 일어날 조건부 확률은

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

로 정의된다.

조건부 확률의 예제1

- 두 개의 동전을 던지는 실험
  - ✓ 표본공간 $S = \{HH, HT, TH, TT\}$, 표본공간의 각 원소는 동일한 확률을 갖는다고 하자.
  - ✓ 사건 A : 첫 번째 동전이 앞면(H)인 사건

    $A = \{HT, HH\}$, $P(A) : \frac{1}{4}$
  - ✓ 사건 B : 두 개의 동전이 모두 앞면(H)인 사건

    $B = \{HH\}$, $P(B) : \frac{1}{4}$
  - ✓ 첫 번째 동전이 앞면(H)라는 조건하에서 두 개의 동전이 모두 앞면(H)일 조건부 확률

    $$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}$$
  - ✓ 사건 C : 최소한 하나의 동전이 앞면(H)인 사건

    $P(C) : ?$
  - ✓ 최소한 하나의 동전이 앞면이라는 조건하에서 둘 다 앞면일 조건부 확률

    $P(B|C) : ?$

Multiplication rule

사건 $A_1, A_2, \ldots, A_n$이 표본공간 S상에 정의되어 있으면, 다음의 공식이 성립한다.

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

Multiplication rule 예제

- 주머니 속에 빨간 공이 1개, 하얀 공이 9개 들어 있다. 이제 1개의 공을 무작위 복원추출하는 실험을 반복한다고 하자. 4번째 추출한 공이 빨간 공일 확률은?
  - ✓ R_4 : 4번째 추출한 공이 빨간색일 사건
  - ✓ W_i : i번째 $(i = 1,2,3)$ 추출한 공이 하얀색일 사건
  - ✓ $W_1 \cap W_2 \cap W_3 \cap R_4$: 공을 4번 뽑았을 때 1,2,3번째는 하얀색 공, 4번째는 빨간색 공이 나오는 사건

$$P(W_1 \cap W_2 \cap W_3 \cap R_4) = P(W_1)P(W_2|W_1)P(W_3|W_1 \cap W_2)P(R_4|W_1 \cap W_2 \cap W_3) = \left[\frac{9}{10}\right]\left[\frac{8}{9}\right]\left[\frac{7}{8}\right]\left[\frac{1}{7}\right]$$

전확률 공식

전확률 공식은 조건부 확률로부터 유래하며, Bayes 정리에 사용하는 공식이다

사건 $B_1, B_2, \ldots, B_k$는 상호배반이며 $(B_i \cap B_j = \emptyset \ ,\ i \neq j)$, $\bigcup_{i=1}^{k} B_k = S$ 라고 하자. 이때 임의의 사건 A에 대하여,

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

가 성립한다.

전확률 공식 예제

- 어떤 공장에 세 개의 생산라인이 있으며, 이들은 각각 전체 생산품의 50%, 30%, 20%를 만들어내며, 각각 2%, 5%, 10%의 불량품을 생산한다. 이제 어떤 제품을 무작위로 추출하였을 때 불량품일 확률을 구해 보자.
  - ✓ 사건 A : 무작위로 추출된 제품이 50% 생산라인으로부터의 것일 사건
  - ✓ 사건 B : 무작위로 추출된 제품이 30% 생산라인으로부터의 것일 사건
  - ✓ 사건 C : 무작위로 추출된 제품이 20% 생산라인으로부터의 것일 사건
  - ✓ 사건 D : 무작위로 추출한 제품이 불량품일 사건

$$P(D) = P(A \cap D) + P(B \cap D) + P(C \cap D)$$

$$= P(A)P(D|A) + P(B)P(D|B) + P(C)P(D|C)$$

$$= 0.5 \times 0.02 + 0.3 \times 0.05 + 0.2 \times 0.10$$

$$= 0.045$$

베이즈 정리

사건 $B_1, B_2, \ldots, B_k$는 상호배반이며 $(B_i \cap B_j = \emptyset\ ,\ i \neq j)$, $\bigcup_{i=1}^{k} B_k = S$ 라고 하자. 이때 사건 A가 일어났다는 조건하에서 사건 B_j가 일어날 확률은

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

로 주어진다.

베이즈 정리 예제 1

- 어떤 질병 A에 대한 양성(+)과 음성(-)반응을 나타내는 검사방법이 있다.
  - ✓ 실제로 질병이 있을 경우 양성반응을 나타낼 확률 :

    $P(+|A) = 0.95$
  - ✓ 질병이 없을 경우에 양성반응을 나타낼 확률 :

    $P(+|A^c) = 0.10$
  - ✓ 어떤 집단에서 무작위추출한 사람이 질병을 가지고 있을 확률 :

    $P(A) = 0.01$
  - ✓ 어떤 사람이 검사반응에서 양성(+)을 나타냈을 때, 그 사람이 실제로 질병 A를 가졌을 확률 :

$$P(A|+) = \frac{P(A \cap +)}{P(+)} = \frac{P(+|A)P(A)}{P(+|A)P(A) + P(+|A^c)P(A^c)} = \frac{0.95 \times 0.01}{0.95 \times 0.01 + 0.10 \times 0.99} = 0.088$$

베이즈 정리 예제 1

- 어떠한 질병이 있는지 없는지 검사를 받을 때 피검자가 관심을 가지는 확률은 질병이 있다는 가정하에서 양성반응을 보이는 확률이나 질병이 없다는 가정하에서 양성반응을 보이는 확률이 아니다. 이 확률은 검사 자체의 신뢰도와 관련된다.

- 피검자가 관심을 가지는 확률은 검사결과가 양성으로 나왔을 경우 또는 검사결과가 음성으로 나왔을 경우 실제로 질병이 존재할 확률, 즉 $P(A|+)$ 또는 $P(A|-)$가 된다.

- 이런 확률은 주어진 정보와 베이즈 정리를 가지고 구할 수 있다.

- $P(+|A) = 0.95$이고 $P(A|+) = 0.088$ 인 이유는 $P(A) = 0.01$ 즉 병을 가지고 있는 사람자체가 적기 때문이다. $P(A|-) = 0.00056$에 비하면 훨씬 높은 수치이다.

베이즈 정리 예제 2

- 각각 1,2,3으로 번호가 붙은 세 개의 주머니가 있는데, 그 안에 5개의 공이 들어 있다. 주머니 i에는 i개의 하얀 공과$(5-i)$개의 검은 공이 들어 있다. 주머니 하나를 무작위로 뽑고, 뽑힌 주머니 안에서 한 개의 공을 무작위로 꺼내는 실험을 하자.

✓ 사건 A : 하얀 공이 뽑히는 사건

$$P(A) = \sum_{i=1}^{3} P(B_i)\,P(A|B_i) = \sum_{i=1}^{3} \left[\frac{1}{3}\right]\left[\frac{i}{5}\right] = \frac{6}{15}$$

✓ 사건 B_i : 주머니 i가 뽑히는 사건

✓ 하얀공이 뽑혔다는 조건하에 뽑힌 하얀 공이 3번 주머니에서 나왔을 확률 :

$$P(B_3|A) = \frac{P(B_3)P(A|B_3)}{\sum_{i=1}^{3} P(B_i)\,P(A|B_i)} = \frac{\frac{1}{3}\times\frac{3}{5}}{\frac{6}{15}} = \frac{1}{2}$$

오즈(odds)

새로운 증거가 나타날 때 **어떤 가설에 대한 확률의 변화**는 이 가설의 오즈 또는 승산에 있어서의 변화로 표현될 수 있다.

사건 A의 오즈는 다음과 같이 표현된다.

$$\frac{P(A)}{P(A^c)} = \frac{P(A)}{1 - P(A)}$$

즉, 사건 A의 오즈는 사건 A가 발생할 가능성이 사건 A가 발생하지 않을 가능성 보다 얼마나 큰지를 설명해 준다.

새로운 증거 B가 나타난 후 새로운 오즈는 다음과 같이 표현된다.

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

오즈(odds) 예제

- 동전 A는 던질 때 앞면이 나타날 확률이 $\frac{1}{4}$인 반면, 동전 B는 던질 때 앞면이 나타날 확률이 $\frac{3}{4}$이다. 이 동전 중 하나를 무작위로 선택하여 두 번 던진다고 하자. 만일 두 번 모두 앞면이 나타났다면, 동전 B가 선택되었을 확률은 얼마인가?
  - ✓ 사건 A : 두 번 모두 앞면이 나올 사건
  - ✓ 사건 B : 동전 B를 택해 던질 사건

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

    $$\frac{P(B|A)}{P(B^c|A)} = \frac{P(B)P(A|B)}{P(B^c)P(A|B^c)} = \frac{\frac{1}{2} \times \left(\frac{3}{4} \times \frac{3}{4}\right)}{\frac{1}{2} \times \left(\frac{1}{4} \times \frac{1}{4}\right)} = 9$$
  - ✓ 승산은 9:1이다.
  - ✓ 즉, 동전 B가 선택되어 던져졌을 확률은 $\frac{9}{10}$이다.

독립사건

① 두 사건 A와 B가 $\boldsymbol{P(A \cap B) = P(A) \cdot P(B)}$를 만족시키면 서로 **독립**(independent)이라고 한다.

② n개의 사건 $A_1, A_2, \ldots, A_n$에 대하여 $\boldsymbol{P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1) \cdot P(A_2) \cdot \cdots \cdot P(A_n)}$가 성립하면 $A_1, A_2, \ldots, A_n$은 **독립**(mutually independent)이라고 한다.

두 사건 A와 B가 독립이면, A의 발생은 B의 발생과 상관 관계가 없다.

독립사건 예제

- 동전을 3회 던지는 실험이 있다고 가정하자.
  - ✓ 사건 H_1 : 첫 번째 던진 결과가 앞면(H)일 사건

    $$P(H_1) = \frac{1}{2}$$
  - ✓ 사건 T_2 : 두 번째 던진 결과가 뒷면(T)일 사건

    $$P(T_2) = \frac{1}{2}$$
  - ✓ 사건 H_3 : 세 번째 던진 결과가 앞면(H)일 사건

    $$P(H_3) = \frac{1}{2}$$
  - ✓ 사건 H_1, T_2의 독립성

    $P(H_1 \cap T_2) = P(HTH, HTT) = \frac{2}{8} = \frac{1}{2} \times \frac{1}{2} = P(H_1) \times P(T_2)$ 즉 독립이다.
  - ✓ 사건 H_1, T_2, H_3의 독립성

    $P(H_1 \cap T_2 \cap H_3) = P(HTH) = \frac{1}{8} = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = P(H_1)P(T_2)P(H_3)$ 즉 독립이다.

독립사건 예제 2

- 어떤 상자에 0과 1로 만들어진 세 자리 숫자가 써진 8장의 카드가 아래와 같이 들어 있다.

1 1 1	1 0 0	0 1 0	0 0 1
1 1 1	1 0 0	0 1 0	0 0 1

✓ 사건 A : 카드 1장을 무작위 추출하였을 때, 카드의 첫 번째 숫자가 1인 사건

✓ 사건 B : 카드 1장을 무작위 추출하였을 때, 카드의 두 번째 숫자가 1인 사건

✓ 사건 C : 카드 1장을 무작위 추출하였을 때, 카드의 세 번째 숫자가 1인 사건

▶ $P(A) = P(B) = P(C) = \frac{4}{8} = \frac{1}{2}$

▶ $P(A \cap B) = P(B \cap C) = P(A \cap C) = \frac{2}{8} = \frac{1}{4}$ □ A, B / B, C / C, A 각각의 쌍은 독립사건을 이룬다.

▶ $P(A \cap B \cap C) = \frac{2}{8} = \frac{1}{4} \neq \frac{1}{8} = P(A)P(B)P(C)$ □ A, B, C는 독립이 아니다.

조건부 독립의 개념

- Rain(비)와 Thunder(천둥)과 Lightning(번개)의 관계를 생각해보자.

  번개가 치는 것을 알게 된다면, 천둥에 대한 추가적인 정보(뒤이어 천둥소리가 들릴 지, 안 들릴 지)는 비가 내린다는 사실에 의해 설명되지 않는다. 왜냐하면 번개가 치면 뒤이어 천둥소리가 들릴 것을 우리는 알기 때문이다.

  ✓ p(천둥 | 비, 번개)

  : 번개가 치고 비가 온다는 사실을 알았을 때 뒤이어 천둥소리가 들릴 확률

  ✓ p(천둥 | 번개)

  : 번개가 친다는 사실을 알았을 때 뒤이어 천둥소리가 들릴 확률

  즉, 두 개의 확률이 같다. 즉, p(천둥 | 비, 번개) = p(천둥 | 번개)

  <u>번개가 치는 정보가 주어졌기 때문에 뒤이어 천둥소리가 들릴 확률은 번개에 영향을 받게 된다.</u> 한마디로 번개의 정보가 있을 때, <u>비의 정보는 필요 없다</u>.

  주의할 것! 일반적으로 비와 천둥은 의존성(dependency)가 있다. 하지만 번개의 정보가 주어졌을 때는 비와 천둥은 조건적으로 독립(conditional dependence)인 것이다.

조건부 독립

- 사건 Z가 주어졌을 때, 사건 X가, 사건 Y에 대하여 조건부 독립이 되기 위한 조건

  $$\boldsymbol{P(X|Y \cap Z) = P(X|Z)}$$

- $P(X|Y \cap Z) = P(X|Z) \Leftrightarrow \boldsymbol{P(X \cap Y|Z) = P(X|Z)P(Y|Z)}$

  ✓ Why?

  $$P(X \cap Y|Z) = \frac{P(X \cap Y \cap Z)}{P(Z)} = \frac{P(X|Y \cap Z)P(Y \cap Z)}{P(Z)} = P(X|Y \cap Z)P(Y|Z) = P(X|Z)P(Y|Z)$$

fast campus

조건부 독립의 사용 예제

"You love to buy some Viagra."라는 메일이 있고, 이 메일은 Spam 메일이다.

Spam 메일을 분류하기 위하여 Naïve Bayesian 분류기를 사용한다고 가정하자.

Naïve Bayesian 분류기는 조건부 독립을 가정한다.

따라서, 다음이 성립한다. $P(\text{buy} \mid \text{to}, \underline{\text{Spam}}) = P(\text{buy} \mid \underline{\text{Spam}})$

Spam이라는 조건이 존재하면 buy가 문장 내에서 보이는 확률은 to에 독립적인 것이다.

즉, 조건부 독립 가정 하에서 다음은 같은 문장이 된다.

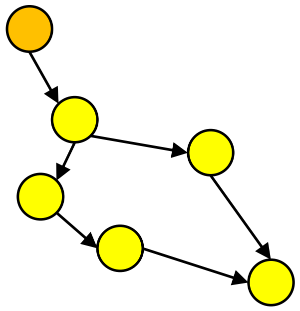
You love to buy some Viagra
You love buy to some Viagra

하지만 영어의 문법 구조상 buy앞에 to가 와야 한다. 문장 내에서 buy는 to에 의존적이다.

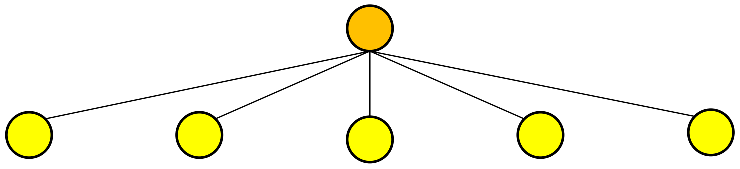
Naïve스러운, 조건부 독립 가정이지만, 이 분류기의 성능은 뛰어나다.

조건부 독립의 사용 예제

Bayesian Network



Naïve Bayesian Network



경우의 수

- 확률의 종류 중 등확률을 고려
  - 유한개의 원소로 이루어진 표본공간을 형성하는 실험들의 모든 가능한 원소가 발생할 확률이 동일하다고 가정

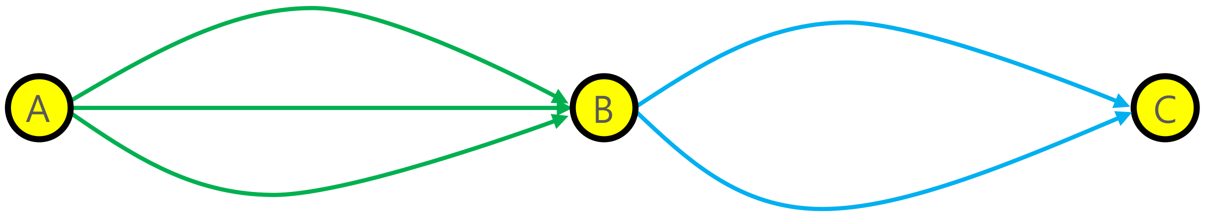
- 어떤 사건 A가 일어날 확률 $P(A)$를 정의
  - $n(A)$: 사건 A가 발생할 수 있는 경우의 수
  - N : 모든 가능한 경우의 수
  - $P(A)$: $\frac{n(A)}{N}$

곱의 법칙

- 사건 A가 일어나는 방법이 m가지이고, 그 각각에 대하여 사건 B가 일어나는 방법이 n가지이면, 사건 A와 B가 동시에 일어나는 방법은 $m \times n$이다.

ex1) 주사위를 1번 던진 후, 동전을 3번 던졌을 때 나올 수 있는 결과의 경우의 수는?

ex2) A에서 C에 도달할 수 있는 총 경로의 수는?



순열

- 서로 다른 n개의 원소 중에서 r개를 선택하여 '순서 있게(ordered)' 늘어놓은 것을 n개에서 r개 택한 순열(permutation)이라고 하고, 늘어놓는 경우의 수를 ${}_nP_r$로 표현한다.

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

ex1) a, b, c, d, e를 일렬로 나열하는 경우의 수는?

ex2) a, a, c, d, e를 일렬로 나열하는 경우의 수는?

조합

- 서로 다른 n개의 원소 중에서 '순서에 관계없이' r개를 선택할 때 이 r개로 이루어진 각각의 집합을 n개에서 r개 택한 조합(combination)이라고 하고, 선택하는 조합의 수를 ${}_nC_r$ 또는 $\binom{n}{r}$로 표현한다.

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

ex1) 한 반의 학생 수가 30명일 때 다음을 구하여라.
 - 반장 1명, 부반장 1명을 뽑는 경우의 수 (순열? 조합?)
 - 주번 2명을 뽑는 경우의 수 (순열? 조합?)

ex2) 라면을 끓여 먹으려고 한다. 떡, 달걀, 치즈, 만두, 파, 김치의 6가지 재료 중
 - 4가지를 선택해서 라면을 끓이는 경우의 수는?
 - 달걀, 치즈는 반드시 포함하여 4가지를 선택해서 라면을 끓이는 경우의 수는?

경우의 수 예제 1

- n개의 공을 n개의 바구니에 임의로 넣을 때 모든 바구니가 한 개씩의 공을 갖게 될 확률 P
  - ✓ n개의 공을 n개의 바구니에 임의로 넣는 경우의 수 :
    n^n
  - ✓ 모든 바구니가 한 개씩의 공을 갖게 될 경우의 수 :
    $n \times (n-1) \times \cdots \times 1 = n!$
  - ✓ $P = \frac{n^n}{n!}$

- k개의 공을 $n(n > k)$개의 바구니에 임의로 넣을 때 모든 바구니가 한 개 이하의 공을 갖게 될 확률 P
  - ✓ k개의 공을 n개의 바구니에 임의로 넣는 경우의 수 :
    n^k
  - ✓ 모든 바구니가 한 개 이하의 공을 갖게 될 경우의 수 :
    $n \times (n-1) \times \cdots \times (n-k+1) = {}_nP_r$
  - ✓ $P = \frac{{}_nP_r}{n^k}$

경우의 수 예제 2

- 어떤 주머니에 하얀 공이 5개 검은 공이 10개 들어 있다. 이 주머니에서 3개의 공을 무작위로 추출하였을 때, 하얀 공이 2개, 검은 공이 1개 뽑힐 확률 P를 구해 보자.

  - 순서를 고려하고 + 복원추출(sampling with replacement)
    - 추출순서를 고려하여 전체 가능한 가짓수 : 15^3
    - 하얀 공(W)이 2개 뽑히는 가짓수 : 5^2
    - 검은 공(B)이 1개 뽑히는 가짓수 : 10^1
    - 하얀 공이 2개, 검은 공이 1개 뽑히는 사건 : WWB , WBW , BWW

P : $\frac{3\times5^2\times10^1}{15^3}$

경우의 수 예제 2

- 어떤 주머니에 하얀 공이 5개 검은 공이 10개 들어 있다. 이 주머니에서 3개의 공을 무작위로 추출하였을 때, 하얀 공이 2개, 검은 공이 1개 뽑힐 확률 P를 구해 보자.

  - 순서를 고려하지 않고 + 비복원추출(sampling with no replacement)
    - 추출순서를 고려하지 않은 전체 가능한 가짓수 : $\binom{15}{3}$
    - 하얀 공(W)이 2개 뽑히는 가짓수 :$\binom{5}{2}$
    - 검은 공(B)이 1개 뽑히는 가짓수 : $\binom{10}{1}$

P : $\frac{\binom{5}{2}\times\binom{10}{1}}{\binom{15}{3}}$

경우의 수 예제 3

- 52개의 카드로 구성된 한 벌의 카드묶음으로 포커게임을 한다. 어떤 사람이 5장의 카드를 받았을 때, 이 패가 스트레이트일 확률 P를 구해 보자.
  - 5장 모두가 같은 그림이면서 숫자가 연속인 스트레이트 플러시는 제외한다.
  - 가능한 숫자조합은 $(A, 2,3,4,5)$, $(2,3,4,5,6)$,…, $(10, J, Q, K, A)$까지 될 수 있다.
  - 가능한 숫자조합 1가지를 뽑는 경우의 수 – 모두 같은 그림을 뽑는 경우의 수 :
    $4^5 - 4$
  - 가능한 숫자조합의 경우의 수 :
    10
  - 스트레이트가 되는 모든 경우의 수 :
    $10(4^5 - 4)$
  - 전체 가능한 경우의 수 :
    $$\binom{52}{5}$$

P : $\frac{10(4^5-4)}{\binom{52}{5}} = 0.0039$

다항계수

n개의 서로 다른 대상물을 크기가 각각 $n_1, n_2, \dots, n_r$인 r개의 그룹으로 분할하는 가능한 방법의 개수. 다항계수 또는 multinomial coefficient라고 한다.

$n_1 + n_1 + \cdots + n_r = n$ 일 때,

$$\binom{n}{n_1, n_2, \dots, n_r} = \frac{n!}{n_1!\, n_2! \dots n_r!}$$