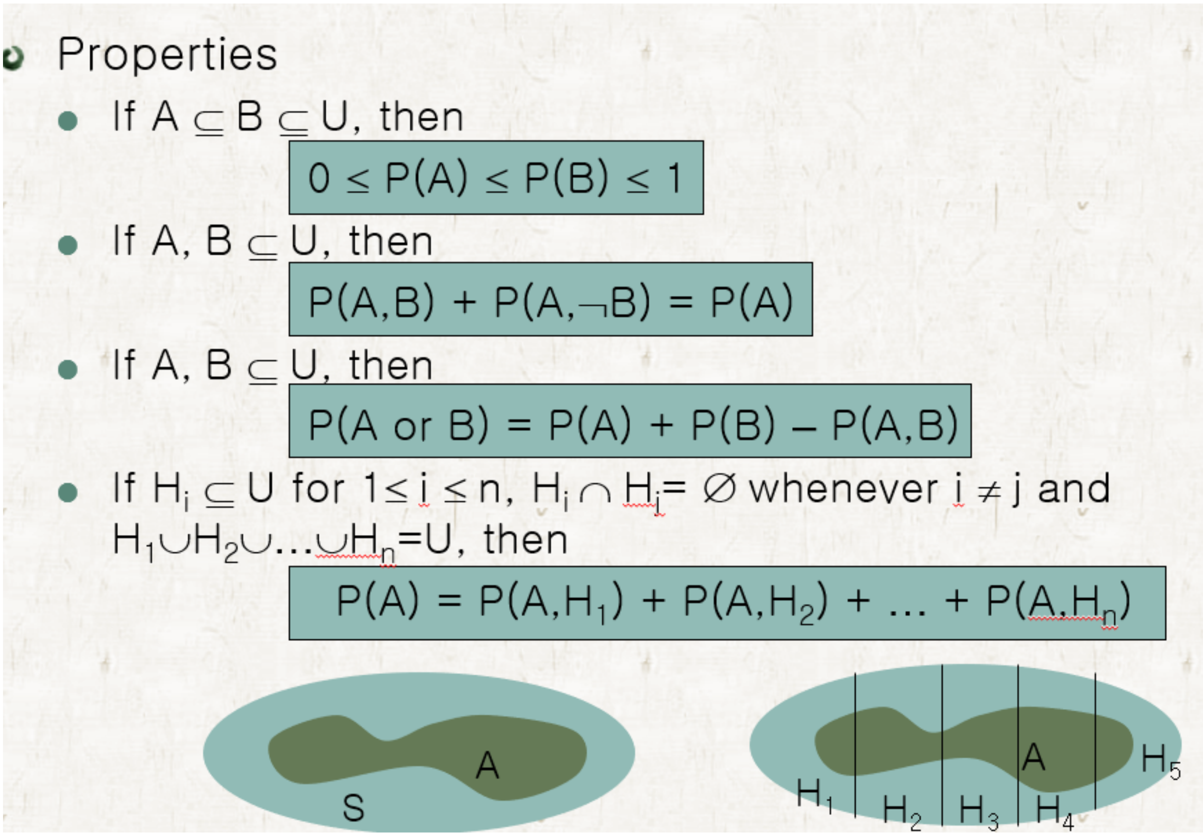
Probability

Probability

- Data analysis is strongly related to probability
- Because of its firm mathematical foundation, many ML approaches have been developed based on probability.
- Naive Bayesian Model, Bayesian Network, Hidden Markov Model,etc..

properties



Conditional Probability

Definition

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

Meaning

- The probability that A will happen under the **assumption that B already happened**
- The probability of A under the **assumption that B is the universal set.**

Properties

- Variations

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

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

- If $A \subseteq B \subseteq U$, then

$$\boxed{0 \leq P(A|C) \leq P(B|C) \leq 1}$$

- If $A, B \subseteq U$, then

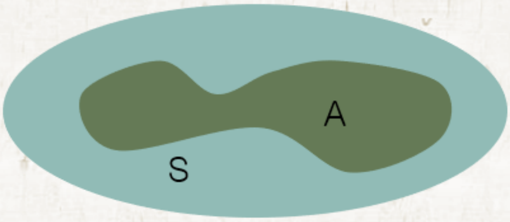
$$\boxed{P(A,B|C) + P(A,\neg B|C) = P(A|C)}$$

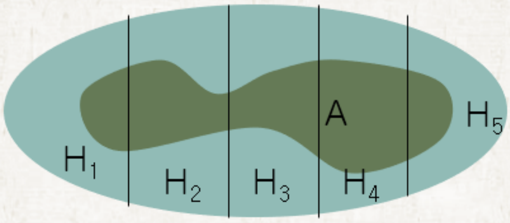
- If $A, B \subseteq U$, then

$$\boxed{P(A \text{ or } B|C) = P(A|C) + P(B|C) - P(A,B|C)}$$

- If $H_i \subseteq S$ for $1 \leq i \leq n$, $H_i \cap H_j = \emptyset$ whenever $i \neq j$ and $H_1 \cup H_2 \cup \ldots \cup H_n = S$, then

$$\begin{aligned} P(A) &= P(A,H_1) + P(A,H_2) + \ldots + P(A,H_n) \\ &= P(A|H_1)P(H_1) + P(A|H_2)\,P(H_2) + \ldots + P(A|H_n)\,P(H_n) \end{aligned}$$





Chaining

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

$$P(A,B,C,D) = P(A|B,C,D)*P(B,C,D)$$
$$P(B,C,D) = P(B|C,D)*P(C,D)$$
$$P(C,D) = P(C|D)*P(D)$$

Order is not important

$$\begin{aligned} P(A,B,C,D) &= P(A|B,C,D)P(B|C,D)P(C|D)P(D) \\ &= P(B|C,A,D)P(C|A,D)P(C|A)P(D) \\ &= P(A|D,C,B)P(D|C,B)P(D|C)P(B) \\ &\ldots \end{aligned}$$

$$P(A,B,C,D) = P(B,C,A,D) = P(A,D,C,B) = \ldots$$

- Bayesian Rule
  - For B, $A \subseteq S$

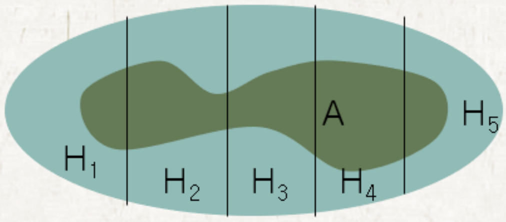
Likelihood of B given A: L(B|A)

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

  - Why important?
    - Usually, prior probabilities are easy to obtain.
    - Usually, one of P(A|B) or P(B|A) are easy to obtain.

=> Evaluating probability hard to obtain from ones easy to obtain

- Bayesian Reasoning
  - If $H_1 \cup H_2 \cup \ldots \cup H_n = S$ and $H_i \cap H_j = \emptyset$ whenever $i \neq j$



$$P(H_i|A) = \frac{P(A|H_i)P(H_i)}{P(A|H_1)P(H_1) + P(A|H_2)\,P(H_2) + \ldots + P(A|H_n)\,P(H_n)}$$

$$= \left(\sum_{j=1}^{n} \frac{P(A|H_j)P(H_j)}{P(A|H_i)P(H_i)} \right)^{-1}$$

Prior probability vs Posterior probability

- Prior probability
  - p(event)
  - Unconditioned probability of an event
  - Probability of an event prior any new evidence
    - Probability of an event without any information(knowledge)
- Posterior probability
  - P(event|evidence)
  - Conditional probability of an event
  - Probability of an event given some new evidence
    - Probability of an event with some piece of information(knowledge)

Independence

Definition

A가 일어나건 안일어나건 B가 일어날 확률에 영향을 미치지 않을때

- A is independent from B if

$$P(A,B) = P(A)P(B)$$

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

- A and C are mutually independent **given B** if

$$P(A,C|B) = P(A|B)P(C|B)$$

$$P(A|B,C) = P(A|B)$$

독립사건을 벤다이어그램으로 표현하면