Unit 2: Conditioning and Independence

MITx Statistics and Data Science

Seung Hyun Sung

2022 10 6

Note

Bayes' rule is the foundation for the field of inference. It guides on how to process data and make inferences about unobserved quantities or phenomena

[1] Conditioning and Bayes'Rule

Conditional Probabilities

- $P(A|B) = \frac{P(A \cap B)}{P(B)}$, defined only when $P(B) > 0$.
- $P(A \cap B|A) = \frac{P(A \cap B)}{P(A)}$
- Properties:
  - $P(A|B) \geq P(B)$, assuming $P(B) > 0$
  - If $A \cap B =$, then $P(A \cup C|B) = P(A|B) + P(C|B)$

Three important tools

- Multiplication rule
  $P(A \cap B) = P(B)P(A|B) + P(A)P(B|A)$
  $P(A^c \cap B \cap C^c) = P(A^c)P(B|A^c)P(C^c|A^c \cap B)$
  $P(A_1 \cap A_2 \cap A_3) = P(A_1)\prod_{i=2}^{n} P(A_i|A_1 \cap ... \cap A_{i_1})$
- Total probability theorem
  Mutually exclusive ($B_i \cap B_j = \oslash$), $\{B_1, B_2, ..., B_k\}$
  $P(B) = \sum_{i=1}^{k} P(A_i)P(B|A_i)$, which is a weighted average of $P(B|A)$. Note that $\sum_i P(A_i) = 1$
- Bayes' rule
  $P(A_i|B) = \frac{A_i \cap B}{P(B)} = \frac{P(A_i)P(B|A_i)}{\sum_{j=1}^{k} P(A_j)P(B|A_j)} = \frac{P(A_i)P(B|A_i)}{P(B)}$
  Inference
  - initial belief $P(A_i)$ on possible causes of an observed event B
    $A_i \xrightarrow{\text{model } P(B|A_i)} B$
  - draw conclusions about causes
    $B \xrightarrow{\text{inference } P(A_i|B)} A_i$

[2] Independence

- Independence of two events
  - Definition of independence: $P(A \cap B) = P(A) \cdot P(B)$.
  - If $P(B|A) = P(B)$, occurrence of A provides no new information about B.
  - If A and B are independence, then A and B^c are independent.

- Conditional independence
  - Given C, the conditional independence is defined as independence under the probability law $P(\cdot|C)$.

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

- Independence does not imply conditional independence.

  In the case below, A and B have no intersection in the condition of C. If A happens, B won't happen in the condition of C. This means that A and B are not independent.

- Independence of a collection of events
  - Event $A_1, A_2, ..., A_n$ are independent if $P(A_i \cap A_j \cap ... \cap A_m) = P(A_i)P(A_j)...P(A_m)$, for any distinct indices $i, j, ..., m$.

- Pairwise independence
  - Independent events must be pairwise independent, but the reverse may not be true.
  - E.g. two independent fair coin tosses. C: the two tosses had the same result. H_1: first toss is H, H_2: second toss is H, $P(H_1) = P(H_2) = 1/2, P(C) = 1/2$.
  - $P(H_1 \cap H_2) = 1/4 = P(H_1)P(H_2)$
  - $P(H_1 \cap C) = 1/4 = P(H_1)P(C)$
  - But, $P(C|H_1 \cap H_2) = 1 \neq P(C) = 1/2$

  So H_1, H_2, C are pairwise independent, but not independent.

- Reliability
  - p_i: probability that unit i is "up" ; u_i: ith unit up, u_i are independent; F_i: ith unit down, F_i are independent.

$$\begin{aligned} P(\text{system up}) &= P(u_1 \cap u_2 \cap u_3) \\ &= P(u_1) \cap P(u_2) \cap P(u_3) \\ &= p_1 p_2 p_3 \end{aligned}$$

$$\begin{aligned} P(\text{system up}) &= P(u_1 \cup u_2 \cup u_3) \\ &= 1 - P(F_1 \cap F_2 \cap F_3) \\ &= 1 - P(F_1)P(F_2)P(F_3) \\ &= 1 - (1 - p_1)(1 - p_2)(1 - p_3) \end{aligned}$$

- In general, if a **serial** sub-system contains m components with success probabilities $p_1, p_2...p_m$, then the probability of success of the entire sub-system is given by

$$P(\text{whole system secceeds}) = p_1 p_2 ... p_m$$

If a parallel sub-system contains m components with success probabilities $p_1, p_2 ... p_m$, then the probability of success of the entire sub-system is given by

$$P(\text{whole system succeeds}) = 1 - (1 - p_1)(1 - p_2)...(1 - p_m)$$