2-3 Conditional Probability

Definition
The *conditional probability* of an event A assuming another event M is

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

for $P(M) > 0$.

Properties

- $M \subseteq A \Rightarrow P(A|M) = 1$.
- $A \subseteq M \Rightarrow P(A|M) = \dfrac{P(A)}{P(M)}$.
- Conditional probabilities are probabilities: $P(A|M)$ satisfies the Axioms of Probability:
    1. $P(A|M) \geq 0$
    2. $P(\mathcal{S}|M) = 1$
    3. $A \cap B = \varnothing \quad \Rightarrow \quad P(A \cup B|M) = P(A|M) + P(B|M)$

Bayes Stuff for events A and B.

- Bayes Rule: $P(A|B) = \dfrac{P(B|A)P(A)}{P(B)}$
- Total Probability Theorem: Let $A_1, \ldots, A_n$ be a partition of $\mathcal{S}$. Then

$$P(B) = P(B|A_1)P(A_1) + \cdots + P(B|A_n)P(A_n).$$

- Bayes Theorem for the partition $A_1, \ldots, A_n$ of $\mathcal{S}$.

$$\begin{aligned} P(A_i|B) &= \frac{P(B|A_i)P(A_i)}{P(B)} \\ &= \frac{P(B|A_i)P(A_i)}{P(B|A_1)P(A_1) + \cdots + P(B|A_n)P(A_n)} \end{aligned}$$

Thomas Bayes (c. 1701 – 7 April 1761)
Presbyterian minister, statistician, and philosopher

Latinisms

Bayes' Rule $P(A|B) = \dfrac{P(B|A)P(A)}{P(B)}$

- $P(A|B)$ is called the *a posteriori* probability.
  - *a posteriori* is a Latin phrase that means "from the later."
  - *a posteriori knowledge* describes the knowledge obtained through experience or observation.
  - *a posteriori reasoning* describes the reasoning based on known facts or past events.
  - *a posteriori reasoning* draws general conclusions from specific (or particular) circumstances—the reasoning is from particular to general. An example is how court cases are argued: inferring the interior disposition of malice from the act of murder.
  - *a posteriori probability* quantifies the uncertainty about the event A given the fact observation that event B has occurred.
- $P(A)$ is called the *a priori* probability.
  - *a priori* is a Latin phrase that means "from the earlier."
  - *a priori knowledge* is independent from any experience and observation. It derives from pure reason.
  - *a priori reasoning* draws particular conclusions from general principles—the reasoning is from from general to specific. A familiar example is mathematical proofs where is a result is deduced from a principle.
  - *a priori probability* quantifies the uncertainty about the event A before any observations are made.
- $P(B|A)$ is called the *likelihood*.
- $P(B)$ does not have a name. The Total Probability Theorem in the context of Bayes' Theorem teaches that $P(B)$ may be expressed in terms of the *likelihood*.

Definition
Two events A and B are called *independent* if

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

Properties

- A and B are independent $\Rightarrow$ A and $\overline{B}$ are independent.
- A and B are independent $\Rightarrow$ $\overline{A}$ and $\overline{B}$ are independent.
- A and B are independent $\Rightarrow$ $P(A|B) = P(A)$.

Definition
Three events A_1, A_2, and A_3 are called (*mutually*) *independent* if

$$P(A_i \cap A_j) = P(A_i)P(A_j), \quad i \neq j$$

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3).$$

Properties

- $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2 \cap A_3)$
- A_1, A_2, A_3 are independent $\Rightarrow$ $P(A_1 \cap A_2 \cap \overline{A}_3) = P(A_1)P(A_2)P(\overline{A}_3)$.

Generalization
The independence of n events can be defined inductively. The events $A_1, \ldots, A_n$ are independent if any $k < n$ of them are independent and

$$P(A_1 \cap \cdots \cap A_n) = P(A_1) \cdots P(A_n)$$

$$P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i).$$

Union Bound

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i).$$