Probability

$A, B$ events such that $p(B) \neq 0$.

$$p(A \mid B) = \frac{p(A \cap B)}{p(B)}$$

Two extremes:

- $p(A \mid B) = 0$, $A$ and $B$ are disjoint events.
- $p(A \mid B) = p(A)$, $A$ and $B$ are independent (we do not gain any information about the probability of $A$ given that $B$ has occurred).

**Def:** $A$ and $B$ are independent if

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

Thus if $B \neq \emptyset$, then $p(A \mid B) = \frac{p(A)p(B)}{p(B)} = p(A)$

**Def:** Events $A_1, A_2, \ldots, A_n$ are mutually independent is $\forall S \subseteq \{1, 2, \ldots, n\}$, $p(\bigcap_{i \in S} A_i) = \prod_{i \in S} p(A_i)$

**Ex:** Pairwise but not mutually independent:

Flip 2 fair coins. $S = \{HH, HT, TH, TT\}$, $p =$ uniform.

- $A =$ “Same outcome” $= \{HH, TT\}$
- $B =$ “First flip is $H$” $= \{HH, HT\}$
- $C =$ “Second flip is $H$” $= \{HH, TH\}$

These events are pairwise independent, but not mutually independent:

- $p(A \cap B) = \frac{1}{4} = p(A)p(B) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$
- $p(B \cap C) = \frac{1}{4} = p(B)p(C) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$
- $p(A \cap C) = p(A)p(C)$
- $p(A \cap B \cap C) = \frac{1}{4} \neq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \Rightarrow$ not mutually independent.

**Ex:** $n$ fair coin flips, $A = HTHHTT$. $p(A) = p(HTHHTT) = \frac{1}{2^n} = p(H)p(T)p(H)\ldots p(T)$.

This first probability (the first $p(H)$) denotes the probability that the first flip is $H$. 

Bayes’ Rule

\[ p(A \mid B) = \frac{p(A \cap B)}{p(B)} \quad \text{and} \quad p(B \mid A) = \frac{p(A \cap B)}{p(A)} \]

By substituting, we get Bayes’ formula:

\[ p(A \mid B) = \frac{p(B \mid A)p(A)}{p(B)} \]

**Theorem:** of Complete Probability

\((S, p)\) where \(S = X_1 \cup X_2 \cup \ldots \cup X_n\) s.t. \(X_i \cap X_j = \emptyset\). For \(E \subseteq S\),

\[ p(E) = \sum_{i=1}^{n} p(E \mid X_i)p(X_i) \]

**Proof.**

\[
\begin{align*}
  p(E) &= p((E \cap X_1) \cup (E \cap X_2) \cup \ldots \cup (E \cap X_n)) \\
        &= p(E \cap X_1) + p(E \cap X_2) + \ldots + p(E \cap X_n) \\
        &= \sum_{i=1}^{n} p(E \mid X_i)p(X_i), \text{ by applying Bayes’ rule}
\end{align*}
\]

Notice that from Bayes’ formula, we can expand out the denominator using a principle called the *law of total probability*:

\[ p(A \mid B) = \frac{p(B \mid A)p(A)}{p(B \mid A)p(A) + p(B \mid \overline{A})p(\overline{A})} \]