Conditional Probability

If two events $A$ and $B$ are not independent we can gain information about $P(A)$ if we know that an event in $B$ has occurred. This is reflected in conditional probability of $A$ given $B$, written as $P(A|B)$:

$$P(A \mid B) = \frac{P(AB)}{P(B)}$$

The unconditional probability $P(A)$ is often called the a priori probability while the conditional probability $P(A|B)$ is often called the a posteriori probability. Note that conditioning may take place in either direction:

$$P(AB) = P(A|B)P(B) = P(B|A)P(A)$$

Conditional probabilities are valid probability measures that satisfy all the fundamental axioms.

If $A$ and $B$ are independent:

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

Example:

$A = \{\text{Algae bloom occurs}\}$
$B = \{\text{Daily average water temperature above 25 deg. C}\}$

Obtain probabilities from long record of daily algae and temperature observations:

Suppose $P(A) = 0.01$, $P(B) = 0.15$, $P(A, B) = 0.005$

Then:

$$P(A \mid B) = \frac{P(AB)}{P(B)} = \frac{0.005}{0.15} = 0.033$$

Probability of a bloom increases significantly if we know that temperature is above 25 deg. C.

Bayes Theorem
Suppose that the sample space \( S \) is divided into a collection of \( n \) mutually exclusive events (sets) called a partition of \( S \):

\[
S = \{A_1, A_2, A_3, \ldots, A_n\}
\]

\( A_i A_j = 0 \quad i \neq j \)

Consider an arbitrary event \( B \) in \( S \), as indicated in the diagram below:

The event \( B \) can be written as the union of the \( n \) disjoint (mutually exclusive) events \( BA_1, BA_2, \ldots, BA_n \):

\[
B = BA_1 + BA_2 + \ldots + BA_n
\]

This implies total probability theorem:

\[
P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \ldots + P(B|A_n)P(A_n)
\]

The total probability theorem and the definition of the conditional probability may be used to derive Bayes theorem:

\[
P(A_i | B) = \frac{P(B | A_i)P(A_i)}{P(B)} = \frac{P(B | A_i)P(A_i)}{P(B | A_1)P(A_1)(B) + \ldots + P(B | A_n)P(A_n)}
\]

Bayes rule updates \( P(A_i) \), given information on the probabilities of obtaining \( B \) when outcomes are \( A_1, A_2, \ldots, A_n \).

Example:

Consider a group of 10 water samples. Exactly 3 are contaminated. Define following events:

<table>
<thead>
<tr>
<th>Event</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( C )</td>
<td>Sample contaminated</td>
</tr>
<tr>
<td>( C' )</td>
<td>Sample not contaminated</td>
</tr>
<tr>
<td>$D$</td>
<td>Contamination detected</td>
</tr>
<tr>
<td>-----</td>
<td>------------------------</td>
</tr>
<tr>
<td>$D'$</td>
<td>Contamination not detected</td>
</tr>
</tbody>
</table>

$P(C) = 0.3$ (based on 3 out of 10 samples contaminated)

Suppose sample analysis technique is imperfect. Based on calibration tests:

$P(D|C) = 0.9$  Successful detection  
$P(D|C') = 0.4$  False alarm

Bayes theorem (replace $A_1$ with $C$, $A_2$ with $C'$, $B$ with $D$):

$$P(C | D) = \frac{P(D | C)P(C)}{P(D | C)P(C) + P(D | C')P(C')} = \frac{(0.9)(0.3)}{(0.9)(0.3) + (0.4)(0.7)} = 0.5$$