AXIOMATIC PROBABILITY AND POINT SETS

The axioms of Kolmogorov. Let S denote an event set with a probability measure P defined over it, such that probability of any event $A \subset S$ is given by $P(A)$. Then, the probability measure obeys the following axioms:

(1) $P(A) \geq 0$,

(2) $P(S) = 1$,

(3) If $\{A_1, A_2, \ldots A_j, \ldots\}$ is a sequence of mutually exclusive events such that $A_i \cap A_j = \emptyset$ for all i, j, then $P(A_1 \cup A_2 \cup \cdots \cup A_j \cup \cdots) = P(A_1) + P(A_2) + \cdots + P(A_j) + \cdots$.

The axioms are supplemented by two definitions:

(4) The conditional probability of A given B is defined by

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

(5) The events A, B are said to be statistically independent if

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

This set of axioms was provided by Kolmogorov in 1936.

The rules of Boolean Algebra. The binary operations of union $\cup$ and intersection $\cap$ are roughly analogous, respectively, to the arithmetic operations of addition $+$ and multiplication $\times$, and they obey a similar set of laws which have the status of axioms:

$$\begin{aligned}
&\text{Commutative law:} \qquad && A \cup B = B \cup A,\\
& && A \cap B = B \cap A,\\
&\text{Associative law:} && (A \cup B) \cup C = A \cup (B \cup C),\\
& && (A \cap B) \cap C = A \cap (B \cap C),\\
&\text{Distributive law:} && A \cap (B \cup C) = (A \cap B) \cup (A \cap C),\\
& && A \cup (B \cap C) = (A \cup B) \cap (A \cup C),\\
&\text{Idempotency law:} && A \cup A = A,\\
& && A \cap A = A.
\end{aligned}$$

De Morgan's Rules concerning complementation are also essential:

$$(A \cup B)^c = A^c \cap B^c \qquad \text{and} \qquad (A \cap B)^c = A^c \cup B^c.$$

Amongst other useful results are those concerning the null set $\emptyset$ and the universl set S:

$$\begin{aligned}
&\text{(i)} \quad A \cup A^c = S, \qquad && \text{(iv)} \quad A \cap S = A,\\
&\text{(ii)} \quad A \cap A^c = \emptyset, && \text{(v)} \quad A \cup \emptyset = A,\\
&\text{(iii)} \quad A \cup S = S, && \text{(vi)} \quad A \cap \emptyset = \emptyset.
\end{aligned}$$

LEMMA: the probability of the complementary event. If A and A^c are complementary events, then

$$P(A^c) = 1 - P(A).$$

Proof. There are

$$A \cup A^c = S \qquad \text{and} \qquad A \cap A^c = \emptyset,$$

Therefore, by Axiom 3,

$$P(A \cup A^c) = P(A) + P(A^c) = 1,$$

since $P(A \cup A^c) = P(S) = 1$, whence $P(A^c) = 1 - P(A)$.

LEMMA: the probability of the null event. The probability of the null event is

$$P(\emptyset) = 0.$$

Proof. Axiom 3 implies that

$$P(S \cup \emptyset) = P(S) + P(\emptyset),$$

since S and $\emptyset$ are disjoint sets by definition, i.e. $S \cap \emptyset = \emptyset$. But also $S \cup \emptyset = S$, so

$$P(S \cup \emptyset) = P(S) = 1,$$

where the second equality is from Axiom 2. Therefore,

$$P(S \cup \emptyset) = P(S) + P(\emptyset) = P(S) = 1,$$

so $P(\emptyset) = 0$.

THEOREM: independence and the complementary event. If A, B are statistically independent such that $P(A \cap B) = P(A)P(B)$, then A, B^c are also statistically independent such that $P(A \cap B^c) = P(A)P(B^c)$.

Proof. Consider

$$A = A \cap (B \cup B^c) = (A \cap B) \cup (A \cap B^c).$$

The final expression denotes the union of disjoint sets, so there is

$$P(A) = P(A \cap B) + P(A \cap B^c).$$

Since, by assumption, there is $P(A \cap B) = P(A)P(B)$, it follows that

$$\begin{aligned} P(A \cap B^c) &= P(A) - P(A \cap B) = P(A) - P(A)P(B) \\ &= P(A)\{1 - P(B)\} = P(A)P(B^c). \end{aligned}$$

THEOREM: the union of of events. The probability that either A or B will happen or that both will happen is the probability of A happening plus the probability of B happening less the probability of the joint occurrence of A and B:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Proof. There is $A \cup (B \cap A^c) = (A \cup B) \cap (A \cup A^c) = A \cup B$, which is to say that $A \cup B$ can be expressed as the union of two disjoint sets. Therefore, according to axiom 3, there is

$$P(A \cup B) = P(A) + P(B \cap A^c).$$

But $B = B \cap (A \cup A^c) = (B \cap A) \cup (B \cap A^c)$ is also the union of two disjoint sets, so there is also

$$P(B) = P(B \cap A) + P(B \cap A^c) \Longrightarrow P(B \cap A^c) = P(B) - P(B \cap A).$$

Substituting the latter expression into the one above gives

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

BAYES' THEOREM

Observe that the formula for conditional probability implies that

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

whence we

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

Consider a set $\Omega = \{H_1, H_2, \ldots, H_n\}$, wherein $H_i \cap H_j = \emptyset; i \neq j$, which comprises all possible explanations of an event E. The evidence of E will cast some light upon the likelihoods of the hypotheses. The *posterior* likelihood of an hypothesis H_i in the light of the event E is

$$P(H_i|E) = \frac{P(E|H_i)P(H_i)}{P(E)}, \quad \text{where}$$

$$P(E) = \sum_i P(E \cap H_i) = \sum_i P(E|H_i)P(H_i).$$

We use the following terminology:

$P(H_i)$ is the prior likelihood of the ith hypothesis H_i,

$P(H_i|E)$ is the posterior likelihood of the ith hypothesis,

$P(E|H_i)$ is the conditional probability of the event E under the hypothesis H_i,

$P(E)$ is the unconditional probability of the event E.

A Bayesian Problem. The dawn train collects any milk churns that were left on the platform of Worplesham station on weekdays. Churns are left on three of the five days. A man arrives at the station not knowing the exact time and thinking that there is a fifty-fifty chance that he has missed the train. Then he notices that there are no milk churns on the platform. How should he reassess the chances that he has missed the train?

Answer: Let T be the event that the train has already passed and let N be the event of there being no milk churns on the platform when I arrive. We have

$$P(T|N) = \frac{P(N|T)P(T)}{P(N)}$$

with

$$P(N) = P(N|T)P(T) + P(N|T^c)P(T^c).$$

We take

$$P(N|T) = 1, \quad P(T) = P(T^c) = \frac{1}{2} \quad \text{and} \quad P(N|T^c) = \frac{2}{5}.$$

Then

$$P(N) = 1.\frac{1}{2} + \frac{2}{5}.\frac{1}{2} = \frac{7}{10}$$

$$\text{and} \quad P(T|N) = 1.\frac{1}{2}.\frac{10}{7} = \frac{5}{7}.$$