2.3: Axioms of Probability

Alex L.

October 10, 2024

Def: We define a **probability function** to be a function which maps a certain event E in a sample space S to a probability of that event happening, $P(E)$, and that obeys the following axioms:

1. $0 \leq P(E) \leq 1$
2. $P(S) = 1$
3. For any disjoint $E_0, E_1, E_2, ...$, $P(\cup E_i) = \sum P(E_i)$, that is, the probability of the union of disjoint events is the sum of the probablity of each individual event.

These are the only three core assumptions that we must make about a probability function, everything else we know about probability can arise from these three axioms.

Deduction: If we combine axioms 2 and 3, we get $P(S \cup \varnothing) = P(S) + P(\varnothing) = 1 + n$. Since the probability function can't output a result greater than one, n must be zero, so we deduce that the probability of the empty set is equal to 0.

Deduction: Notice that an event E and its complement E^C always partition the event space S, so $P(S) = P(E \cup E^C) = 1$, therefore, the probability of something occuring OR it not occuring is always 1.

By restructuring the above deduction, we get that $P(E^C) = 1 - P(E)$.

Theorem: If $E \subset F$ then $P(E) \leq P(F)$.

Proof: If $E \subset F$, then there must be some elements in F that are not in E. Those elements will be in $E^C \cap F$. Then, $F = E + (E^C \cap F)$ so $P(F) = P(E) + P(E^C \cap F)$, and since $P(E^C \cap F) \geq 0$, then $P(F) \geq P(E)$