Probability

Definition: Probability

Probability is a function defined on a space of events that satisfies the following axioms:

1. $\forall\, E \subseteq S, P(E) \geq 0$
2. $P(\mathcal{S}) = 1$
3. Let $\{E_1, E_2, \ldots\}$ be a countably-infinite set of pairwise disjoint events:

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$$

Definition: Relative Frequency

Assume that an experiment is repeated n times and event E occurred $n(E)$ times. The *relative frequency* of E is given by:

$$\frac{n(E)}{n}$$

The probability of E can then be interpreted as:

$$P(E) = \lim_{n->\infty} \frac{n(E)}{n}$$

Theorem

$P(\emptyset) = 0$

Proof. Let $\{E_1, E_2, \ldots\}$ be a countably-infinite set of events such that all the $E_i = \emptyset$. Since $E_i \cap E_j = \emptyset \cap \emptyset = \emptyset$, the E_i are pairwise disjoint. So, from the third axiom:

$$P(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i)$$
$$P(\bigcup_{i=1}^{\infty} \emptyset) = \sum_{i=1}^{\infty} P(\emptyset)$$
$$P(\emptyset) = \sum_{i=1}^{\infty} P(\emptyset)$$

But this can only happen when $P(\emptyset) = 0$. ∎

Theorem

Let $\{E_1, \ldots, E_k\}$ be a finite set of pairwise disjoint events:

$$P\left(\bigcup_{i=1}^{k} E_i\right) = \sum_{i=1}^{k} P(E_i)$$

Proof. Let $\{E_{k+1}, E_{k+2}, \ldots\}$ be a countably-infinite set of events such that all the $E_i = \emptyset$. By the third axiom and previous theorem:

$$P(\bigcup_{i=1}^{k} E_i) = P(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} E_i = \sum_{i=1}^{k} E_i$$

■

Theorem

$P(E) = 1 - P(E^c)$

Proof. By definition, E and E^c are disjoint. So, by the previous theorem:

$$P(E \cup E^c) = P(E) + P(E^c)$$
$$P(\mathcal{S}) = P(E) + P(E^c)$$

But by the second axiom, $P(\mathcal{S}) = 1$, and therefore:

$$P(E) + P(E^c) = 1$$

and:

$$P(E) = 1 - P(E^c)$$

■

Corollary

$0 \leq P(E) \leq 1$

Theorem

$A \subseteq B \implies P(A) \leq P(B)$

Proof. Let $C = B - A$, and thus $B = A \cup C$ and $A \cap C = \emptyset$. By the previous theorem:

$$P(B) = P(A \cup C) = P(A) + P(C)$$

But, by the first axiom, $P(C) \geq 0$. Therefore:

$$P(A) \leq P(B)$$

■

Theorem: PIE

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

Proof. First, decompose B into two disjoint events:

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

Now, decompose $A \cup B$ into two disjoint events:

$$\begin{aligned}
&A \cup B = A \cup (A^c \cap B) \\
&P(A \cup B) = P\left(A \cup (A^c \cap B)\right) = P(A) + P(A^c \cap B) = P(A) + P(B) - P(A \cap B)
\end{aligned}$$

∎

Example

In a large discrete math class, 55% of the students are math majors, 35% are CS majors, and 5% are dual majors (in math and CS). What percentage of the class majors in neither of them?

Let:

$$\begin{aligned}
&A = \{\text{majors in math}\} \\
&B = \{\text{majors in CS}\}
\end{aligned}$$

$$\begin{aligned}
P(A^c \cap B^c) &= P\left((A \cap B)^c\right) \\
&= 1 - P(A \cap B) \\
&= 1 - (P(A) + P(B) - P(A \cap B)) \\
&= 1 - (0.55 + 0.35 - 0.05) \\
&= 1 - 0.985 \\
&= 0.15
\end{aligned}$$

This theorem can be expanded to three events:

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