Axioms of Probability

- **Probability** is a mapping $p : S \to (-\infty, \infty)$ that assigns to each event $E$ in the sample space $S$ a real number $p(E)$ satisfying:
  
  **Axiom 1**: $0 \leq p(E) \leq 1$.
  **Axiom 2**: $P(S) = 1$.
  **Axiom 3**: If $E_1, E_2, E_3, \ldots$ are disjoint, then
  $$p(E_1 \cup E_2 \cup E_3 \cup \ldots) = p(E_1) + p(E_2) + p(E_3) + \ldots.$$ 

- **Fact**: $P(\emptyset) = 0$
  This follows from $S = S \cup \emptyset$ (a disjoint union), since then $P(S) = P(S \cup \emptyset) = P(S) + P(\emptyset)$.

- **Fact**: If $E \subseteq F$, then $P(E) \leq P(F)$.
  This follows from $F = E \cup E^c F$ (a disjoint union), so $P(F) = P(E \cup E^c F) = P(E) + P(E^c F) \geq P(E)$.

- **Fact**: $P(E \cup F) = P(E) + P(F) - P(EF)$.
  This follows from
  
  $$E = EF \cup EF^c$$
  $$F = EF \cup E^c F$$
  $$E \cup F = EF^c \cup E^c F \cup EF$$
  $$P(E \cup F) = P(EF^c) + P(E^c F) + P(EF)$$
  $$= (P(EF^c) + P(EF)) + (P(E^c F) + P(EF)) - P(EF)$$
  $$= P(E) + P(F) - P(EF).$$

- For coin flipping experiments, we write $H$ for Heads and $T$ for Tails. For example, if we flip two coins, then we write the sample space as $S = \{HH, HT, TH, TT\}$. The event that the coin flips are different is $E = \{HT, TH\}$, and its complement is $E^c = \{HH, TT\}$. If the coins are fair and independent, then each element of $S$ has probability $1/4$.

  As another example, if we roll two dice (i.e. 6-sided), then the sample space is
  $$S = \{(1,1), (1,2), \ldots, (1,6), (2,1), (2,2), \ldots, (2,6), \ldots, (6,1), \ldots, (6,6)\}$$
  which has 36 elements in it. Each element of the sample space is a pair of the form $(i,j)$. The event that the dice add up to at least 11 is $E = \{(5,6), (6,5), (6,6)\}$. If the dice are fair and independent, then each element of $S$ has probability $1/36$. 