In probability theory, axioms serve as the foundational rules that govern how probabilities are assigned and manipulated. The consequences of these axioms lead to many important results and properties. The three fundamental axioms of probability (Kolmogorov's axioms) are:

1. **Non-negativity:** P(A)≥0P(A) \geq 0 for any event AA.
2. **Normalization:** P(Ω)=1P(\Omega) = 1, where Ω\Omega is the sample space.
3. **Additivity:** If AA and BB are mutually exclusive (i.e., A∩B=∅A \cap B = \emptyset), then P(A∪B)=P(A)+P(B)P(A \cup B) = P(A) + P(B).

**Key Consequences of These Axioms:**

1. **Probability of the Empty Set (Impossible Event)**
   * Since AA and ∅\emptyset are mutually exclusive, setting A=∅A = \emptyset in the additivity axiom gives: P(∅)+P(A)=P(A)P(\emptyset) + P(A) = P(A) This implies P(∅)=0P(\emptyset) = 0, meaning the probability of an impossible event is always zero.
2. **Monotonicity (Subset Rule)**
   * If A⊆BA \subseteq B, then P(A)≤P(B)P(A) \leq P(B). This follows from the fact that BB can be written as A∪(B−A)A \cup (B - A), and by additivity: P(B)=P(A)+P(B−A)≥P(A)P(B) = P(A) + P(B - A) \geq P(A)
3. **Complement Rule**
   * The probability of an event’s complement is: P(Ac)=1−P(A)P(A^c) = 1 - P(A) This follows from the fact that AA and AcA^c are mutually exclusive and their union is the whole sample space Ω\Omega, so: P(A)+P(Ac)=P(Ω)=1P(A) + P(A^c) = P(\Omega) = 1
4. **Inclusion-Exclusion Principle**
   * For any two events AA and BB, P(A∪B)=P(A)+P(B)−P(A∩B)P(A \cup B) = P(A) + P(B) - P(A \cap B) This accounts for the overlap when adding probabilities of two non-disjoint events.
5. **Union Bound (Boole's Inequality)**
   * For any finite or countable collection of events A1,A2,…A\_1, A\_2, \dots, P(A1∪A2∪… )≤P(A1)+P(A2)+…P(A\_1 \cup A\_2 \cup \dots) \leq P(A\_1) + P(A\_2) + \dots This is a generalization of the additivity property when events are not necessarily disjoint.
6. **Law of Total Probability**
   * If B1,B2,...,BnB\_1, B\_2, ..., B\_n form a partition of the sample space Ω\Omega (i.e., they are mutually exclusive and their union is Ω\Omega), then for any event AA, P(A)=∑i=1nP(A∩Bi)P(A) = \sum\_{i=1}^{n} P(A \cap B\_i) If we use conditional probability, this can be rewritten as: P(A)=∑i=1nP(A∣Bi)P(Bi)P(A) = \sum\_{i=1}^{n} P(A | B\_i) P(B\_i)
7. **Bayes' Theorem**
   * Derived from the definition of conditional probability, Bayes' theorem states: P(Bi∣A)=P(A∣Bi)P(Bi)P(A)P(B\_i | A) = \frac{P(A | B\_i) P(B\_i)}{P(A)} where BiB\_i are elements of a partition of Ω\Omega.