Review of Probability Definitions and Formulas

1. **Definition of a Probability Function on a Sample Space**

Suppose that S is a sample space associated with an experiment. A **probability function** on S is a function P which assigns to each event A in S (that is, to each subset A of S) a real number $P(A)$ and which satisfies the following axioms:

Axiom 1 (*Nonnegativity*): $P(A) \geq 0$ for every event A.

Axiom 2: $P(S) = 1$.

Axiom 3 (*Countable Additivity*): If $A_1, A_2, A_3, \ldots$ are pairwise mutually exclusive events in S (that is, $A_i \cap A_j = \varnothing$ if $i \neq j$), then

$$P\left(A_1 \cup A_2 \cup A_3 \cup \cdots\right) = \sum_{i=1}^{\infty} P\left(A_i\right).$$

2. **Some Properties of Probability Functions**

If P is a probability function on a sample space S, then:

a. $P(\overline{E}) = 1 - P(E)$. [From Axioms 2 & 3, since $S = E \cup \overline{E}$ and $E \cap \overline{E} = \varnothing$.]

b. $P(\varnothing) = 0$. [From (a) and Axiom 2, since $\varnothing = \overline{S}$.]

c. $P(A) = P(A \cap B) + P(A \cap \overline{B})$ and

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$ if A and B are any two events.

These follow from Axiom 3, using

$A = (A \cap \overline{B}) \cup (A \cap B)$,

$B = (\overline{A} \cap B) \cup (A \cap B)$, and

$A \cup B = (A \cap \overline{B}) \cup (A \cap B) \cup (\overline{A} \cap B)$.

d. $P(A \cup B) = P(A) + P(B)$, if A and B are mutually exclusive.

e. $P(A - B) = P(A \cap \overline{B}) = P(A) - P(A \cap B)$ [from part c]

3. **Conditional Probability**

   a. **Definition:** $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$, provided $P(B) \neq 0$.

   b. **Multiplication Rule:** $P(A \cap B) = P(B) \cdot P(A|B)$, provided $P(B) \neq 0$.

   c. **Extended Multiplication Rule:** If $P(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) \neq 0$, then
   $P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)\, P(A_2|A_1)\, P(A_3|A_1 \cap A_2) \cdots P(A_n|\, A_1 \cap A_2 \cap \cdots \cap A_{n-1})$.

   d. $P(A \cup B \mid C) = P(A \mid C) + P(B \mid C) - P(A \cap B \mid C)$ provided $P(C) \neq 0$.

   e. $P(\overline{A} \mid C) = 1 - P(A \mid C)$ provided $P(C) \neq 0$.

4. A and B are **independent** events if and only if $P(A \cap B) = P(A) \cdot P(B)$.

   a. If A and B are independent and $P(B) \neq 0$, then $P(A \mid B) = P(A)$.

   b. If $P(B) \neq 0$ and $P(A \mid B) = P(A)$, then A and B are independent.

   c. If A and B are independent, then

      (i) $\overline{A}$ and B are independent;

      (ii) A and $\overline{B}$ are independent; and

      (iii) $\overline{A}$ and $\overline{B}$ are independent.

5. Events $A_1, A_2, \ldots, A_n$ are **mutually independent** if and only if for every subset $A_{1'}, A_{2'}, \ldots, A_{r'}$, $r \leq n$, of these events

$$P(A_{1'} \cap A_{2'} \cap \cdots \cap A_{r'}) = P(A_{1'})\, P(A_{2'}) \cdots P(A_{r'}).$$

6. **Law of Total Probability.** If $B_1, B_2, \ldots, B_n$ partition the sample space and A is any event, then

$$\begin{aligned} P(A) &= P(A \cap B_1) + P(A \cap B_2) + \cdots + P(A \cap B_n) \\ &= P(B_1) \cdot P(A \mid B_1) + P(B_2) \cdot P(A \mid B_2) + \cdots + P(B_n) \cdot P(A \mid B_n) \end{aligned}$$

7. **Bayes' Theorem:** If $B_1, B_2, \ldots, B_k$ partition S and A is any event, then

$$P(B_j|A) = \frac{P(B_j)P(A|B_j)}{P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + \cdots + P(B_k)P(A|B_k)}, \text{ for } j = 1, 2, \ldots, k.$$