Conditional Probability

Definition and properties

1. **Definition:** The *conditional probability of A given B* is denoted by $P(A|B)$ and defined by the formula
$$P(A|B) = \frac{P(AB)}{P(B)},$$
provided $P(B) > 0$. (If $P(B) = 0$, the conditional probability is not defined.) (Recall that AB is a shorthand notation for the intersection $A \cap B$.)

2. **Rules for conditional probabilities:** The function $P_B(A) = P(A|B)$, considered as a function of A, with the conditioning event B fixed, satisfies the Kolmogorov axioms, and any rules derived from these axioms, in particular:
   - **Addition formula for conditional probabilities:** $P(A_1 \cup A_2 \cup \cdots | B) = P(A_1|B) + P(A_2|B) + \cdots$ for mutually disjoint events $A_1, A_2, \ldots$.
   - **Complement rule for conditional probabilities:** $P(A^c|B) = 1 - P(A|B)$

3. **Multiplication rule (or chain rule) for conditional probabilities:** $P(AB) = P(A|B)P(B)$. More generally, for any events $A_1, \ldots, A_n$,
$$P(A_1A_2 \ldots A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1A_2) \ldots P(A_n|A_1A_2 \ldots A_{n-1}).$$

Notes

- **$P(A|B)$ is not the same as $P(B|A)$:** In contrast to set-theoretic operations like union or intersection, in conditional probabilities the order of the sets matters.

- **$P(A|B^c)$ is not the same as $1 - P(A|B)$:** The complement rule only holds with respect to the first argument.

Interpretations

- **Partial knowledge interpretation:** $P(A|B)$ can be interpreted as the probability that A occurs, *if it is known that B has occurred.* Depending on the particular situation, the additional knowledge about B may make the occurrence of A more likely (i.e., $P(A|B) > P(A)$), less likely (i.e., $P(A|B) < P(A)$), or have no effect (i.e., $P(A|B) = P(A)$). In the latter case, A and B are called *independent.*

- **Area interpretation:** An easy way to visualize conditional probabilities is as *relative areas* in Venn diagrams: $P(A|B)$ represents the percentage of the area of B that is occupied by A.

- **Recognizing conditional probabilities:** In word problems, conditional probabilities can usually be recognized by words like "given", "if", or "among" (e.g., in the context of subpopulations). There are, however, no hard rules, and you have to read the problem carefully and pay attention to the entire context of the problem to determine whether a given probability represents an ordinary probability (e.g., $P(AB)$) or a conditional probability (e.g., $P(A|B)$ or $P(B|A)$).