$$P(\mathscr{A}^{\complement} \cap \mathscr{B}) = P(\mathscr{B}) - P(\mathscr{A} \cap \mathscr{B}) \tag{6.47}$$

Plugging Equation 6.47 into Equation 6.43 we have,

$$\begin{aligned} P(\mathscr{A} \cup \mathscr{B}) &= P(\mathscr{A}) + P(\mathscr{B}) - P(\mathscr{A} \cap \mathscr{B}) \\ &= P(\mathscr{A}) + P(\mathscr{B}) - P(\mathscr{A}, \mathscr{B}) \end{aligned} \tag{6.48}$$

□

Property 6.3 may be extended to three events, $\mathscr{A}_1$, $\mathscr{A}_2$, and $\mathscr{A}_3$ as follows,

$$\begin{aligned} P(\mathscr{A}_1 \cup \mathscr{A}_2 \cup \mathscr{A}_3) = {} & P(\mathscr{A}_1) + P(\mathscr{A}_2) + P(\mathscr{A}_3) - \\ & P(\mathscr{A}_1, \mathscr{A}_2) - P(\mathscr{A}_2, \mathscr{A}_3) - P(\mathscr{A}_1, \mathscr{A}_3) + \\ & P(\mathscr{A}_1, \mathscr{A}_2, \mathscr{A}_3) \end{aligned} \tag{6.49}$$

This can be easily extended to any number of events, but the generalization notation would be somewhat complicated, so it is not shown here. Keep in mind that for a larger number of events, as in the 3-event case, the probabilities of all possible combinations of intersections of events must be subtracted from the sum of the probabilities of all individual events.

Definition 6.36 (Conditional Probability). *If $\mathscr{A} \subset \mathscr{X}, \mathscr{B} \subset \mathscr{X}$, and $P(\mathscr{B}) > 0$, then the probability of event $\mathscr{A}$ given that event $\mathscr{B}$ has occurred is called the conditional probability of $\mathscr{A}$ given $\mathscr{B}$ and is written as,*

$$P(\mathscr{A}|\mathscr{B}) = \frac{P(\mathscr{A}, \mathscr{B})}{P(\mathscr{B})} \tag{6.50}$$

or equivalently,

$$P(\mathscr{A}, \mathscr{B}) = P(\mathscr{A}|\mathscr{B})P(\mathscr{B}) = P(\mathscr{B}|\mathscr{A})P(\mathscr{A}) \tag{6.51}$$

where $P(\mathscr{A}, \mathscr{B})$ (or $P(\mathscr{A} \cap \mathscr{B})$) is called the joint probability of events $\mathscr{A}$ and $\mathscr{B}$. Note that Equation 6.51 does not need the requirements that $P(\mathscr{B}) > 0$ or $P(\mathscr{A}) > 0$.

Consider the two cases where $\mathscr{A} \subset \mathscr{B}$ and $\mathscr{B} \subset \mathscr{A}$. The condition probability, $P(\mathscr{A}|\mathscr{B})$, will have the following properties for each case,

1. $\mathscr{A} \subset \mathscr{B}$,

$$\begin{aligned} P(\mathscr{A}|\mathscr{B}) &= \frac{P(\mathscr{A})}{P(\mathscr{B})} \\ &\geq P(\mathscr{A}) \end{aligned} \tag{6.52}$$

2. $\mathscr{B} \subset \mathscr{A}$,

$$P(\mathscr{A}, \mathscr{B}) = P(\mathscr{B}) \implies P(\mathscr{A}|\mathscr{B}) = 1 \tag{6.53}$$