1. Let $A$ and $B$ be independent events. Show that $A$ and $B^c$ are independent.

**Solution:** We write $A$ as the union of two disjoint sets: $A = (A \cap B^c) \cup (A \cap B)$. Then, by the additivity axiom of probability,

$$P(A \cap B^c) = P(A) - P(A \cap B).$$

Since $A$ and $B$ are independent, $P(A \cap B) = P(A)P(B)$. Therefore,

$$P(A \cap B^c) = P(A) - P(A)P(B) = P(A)(1 - P(B)) = P(A)P(B^c),$$

which implies that $A$ and $B^c$ are independent.

2. Let $A$ and $B$ be events with $P(A) > 0$ and $P(B) > 0$. We say that an event $B$ suggests an event $A$ if $P(A | B) > P(A)$, and does not suggest event $A$ if $P(A | B) < P(A)$. Show that $B$ suggests $A$ if and only if $A$ suggests $B$.

**Solution:** We have $P(A \cap B) = P(A | B)P(B)$, so $B$ suggests $A$ if and only if $P(A \cap B) > P(A)P(B)$, which is equivalent to $A$ suggesting $B$, by symmetry. (Alternatively, one may use Bayes’ theorem to establish the statement.)

Let $X,Y,Z$ be discrete random variables with joint PMF $p$.

3. Show that $p(x | y, z) = p(x, y | z)/p(y | z)$ for all $x, y, z$ with $p(y, z) > 0$.

**Solution:** By the definition of conditional probability,

$$p(x | y, z) = \frac{p(x, y, z)}{p(y, z)}, \quad \forall x, y, z \text{ with } p(y, z) > 0,$$

$$p(x, y | z) = \frac{p(x, y, z)}{p(z)}, \quad \forall x, y, z \text{ with } p(z) > 0,$$

$$p(y | z) = \frac{p(y, z)}{p(z)}, \quad \forall y, z \text{ with } p(z) > 0.$$

Since $p(y, z) > 0$ implies $p(z) > 0$, we have for all $x, y, z$ with $p(y, z) > 0$,

$$p(x | y, z) = \frac{p(x, y, z)}{p(y, z)} = \frac{p(x, y, z)p(z)}{p(y, z)p(z)} = \frac{p(x, y | z)}{p(y | z)}.$$
4*. Show that the following are all equivalent definitions of the conditional independence $X \perp Y \mid Z$:

\begin{align*}
&p(x, y, z) = p(x \mid z)p(y \mid z)p(z), \quad \forall x, y, z. \quad (1) \\
p(x, y, z) = p(x, z)p(y, z)/p(z), \quad \forall x, y, z \text{ with } p(z) > 0. \quad (2) \\
p(x, y, z) \text{ has the form } a(x, z)b(y, z), \quad \forall x, y, z. \quad (3) \\
p(x \mid y, z) = p(x \mid z), \quad \forall x, y, z \text{ with } p(y, z) > 0. \quad (4) \\
p(x \mid y, z) \text{ has the form } a(x, z), \quad \forall x, y, z \text{ with } p(y, z) > 0. \quad (5)
\end{align*}

In the above, $a$ and $b$ are some real-valued functions.

**Solution:** We prove that (1)-(5) are equivalent to each other by showing “(1) $\iff$ (2),” “(1) $\iff$ (4),” “(3) $\Rightarrow$ (4) $\Rightarrow$ (5) $\Rightarrow$ (3),” in this order. Whenever multiplication involves a quantity that appears to be undefined, we define $0 \cdot $ (undefined value) $= 0$ and $c \cdot $ (undefined value) $= $ (undefined value) for $c \neq 0$. Two quantities $c_1$ and $c_2$ are considered to be unequal if both of them are undefined.

(1) $\iff$ (2): Since Eq. (1) always holds when $p(z) = 0$, we only need to show the equivalence of Eq. (1) and Eq. (2) for all $x, y, z$ with $p(z) > 0$. Using the fact $p(x, z) = p(x \mid z)p(z)$ and $p(y, z) = p(y \mid z)p(z)$, we have for all $x, y, z$ with $p(z) > 0$,

$$p(x \mid z)p(y \mid z)p(z) = p(x \mid z)p(y, z) = p(x, z)p(y, z)/p(z),$$

which establishes the claimed equivalence.

(1) $\iff$ (4): When $p(y, z) = 0$, $p(y \mid z)p(z) = 0$ and $p(x, y, z) = 0$. This shows that Eq. (1) always holds when $p(y, z) = 0$, so we only need to prove the equivalence of Eq. (1) and Eq. (4) for all $x, y, z$ with $p(y, z) > 0$. For all such $x, y, z$, using the fact $p(x \mid y, z) = p(x, y, z)/p(y, z)$, we have

$$p(x \mid y, z) = p(x \mid z) \iff p(x, y, z) = p(x \mid z)p(y, z) = p(x, z)p(y, z)/p(z),$$

which establishes the claimed equivalence.

(3) $\Rightarrow$ (4): When (3) holds, we have

$$p(y, z) = b(y, z)\sum_{x'}a(x', z), \quad p(x, z) = a(x, z)\sum_{y'}b(y', z), \quad p(z) = \sum_{x'}a(x', z)\cdot \sum_{y'}b(y', z).$$

For all $x, y, z$ with $p(y, z) > 0$, we have $p(z) > 0$ and

$$p(x \mid y, z) = \frac{p(x, y, z)}{p(y, z)} = \frac{a(x, z)}{\sum_{x'}a(x', z)},$$

$$p(x \mid z) = \frac{p(x, z)}{p(z)} = \frac{a(x, z)\sum_{y'}b(y', z)}{\sum_{x'}a(x', z)\cdot \sum_{y'}b(y', z)} = \frac{a(x, z)}{\sum_{x'}a(x', z)},$$

which implies (4).

(4) $\Rightarrow$ (5) $\Rightarrow$ (3): It is evident that (4) implies (5). Consider now the case where (5) holds. If the function $a(x, z)$ is not defined at some points $(x, z)$, we can always define the values of the function at those points to be zero, say, thereby extending the function $a$ to the entire space of $(x, z)$. Using the identity $p(x, y, z) = p(x \mid y, z)p(y, z)$, we have $p(x, y, z) = a(x, z)p(y, z)$, which implies (3).
Calculations of conditional probabilities:

5. Two fair coins are tossed and at least one coin came up heads. What is the probability that both coins came up heads?

Solution: Let $A$ denote the event that at least one coin come up heads and $B$ the event that both coins come up heads. Then $A = \{HH, HT, TH\}$, $B = \{HH\}$, and $A \cap B = B$. Since the coins are fair, all outcomes are equally likely, so $P(A) = 3/4$, $P(B) = 1/4$, and the desired probability, $P(B|A)$, is

$$P(B|A) = P(A \cap B)/P(A) = \frac{3}{4}/\frac{3}{4} = \frac{1}{3}. \quad \Box$$

6. A test for a certain rare disease is assumed to be correct 95% of the time: if a person has (does not have) the disease, the test results are positive (negative) with probability 0.95. A random person drawn from a certain population has probability 0.001 of having the disease. Given that the person just tested positive, what is the probability that the person has the disease?

Solution: Let $A$ denote the event that the person has the disease and $B$ the event that the test results are positive. From the statement of the problem, we have

$$P(B|A) = P(B^c|A^c) = 0.95, \quad P(A) = 0.001.$$ To calculate the desired probability, $P(A|B)$, we use Bayes’ theorem and the fact

$$P(B) = P(B \cap A) + P(B \cap A^c) = P(B|A)p(A) + P(B|A^c)p(A^c)$$

to obtain

$$P(A|B) = \frac{P(B|A)p(A)}{P(B)} = \frac{P(B|A)p(A)}{P(B|A)p(A) + P(B|A^c)p(A^c)} \approx 0.0187. \quad \Box$$

About Markov chains:

7. A mouse moves along a tiled corridor with $2m$ tiles, where $m > 1$. From each tile $i \neq 1, 2m$, it moves to either tile $i-1$ or $i+1$ with equal probability. From tile 1 or $2m$, it moves to tile 2 or $2m-1$, respectively, with probability 1. Each time the mouse moves to a tile $i \leq m$ or $i > m$, an electronic device outputs a signal $L$ or $R$, respectively. Can the generated sequence of signals $L$ and $R$ be described as a Markov chain with states $L$ and $R$? Suppose now the device outputs $L$ or $R$ only when the mouse moves to tile 1 or $2m$, respectively. Can the generated sequence of $L$ and $R$ be described as a Markov chain with states $L$ and $R$?

Solution: Let $\{X_n\}$ denote the generated sequence of $L$ or $R$. In the case of the first question, $\{X_n\}$ cannot be described as a Markov chain with states $L$ and $R$, because $P(X_{n+1} = L|X_n = R, X_{n-1} = L) = 1/2$, while $P(X_{n+1} = L|X_n = R, X_{n-1} = R, X_{n-1} = L) = 0$.

In the case of the second question, $\{X_n\}$ can be described as a Markov chain with states $L$ and $R$. We argue as follows. The sequence of positions of the mouse can be described as a Markov chain on the state space $\{1, 2, \ldots, 2m\}$. At any time $t$, given that the mouse just moved to tile 1, the probability of any event that concerns only its future positions does not depend on its positions before time $t$. Therefore, for any $x_1, \ldots, x_n \in \{L, R\}$ and $x_{n+1} \in \{L, R\}$,

$$P(X_{n+1} = x_{n+1} | X_n = L, X_{n-1} = x_n, \ldots, X_1 = x_1) = P(X_{n+1} = x_{n+1} | X_n = L),$$

and similarly,

$$P(X_{n+1} = x_{n+1} | X_n = R, X_{n-1} = x_n, \ldots, X_1 = x_1) = P(X_{n+1} = x_{n+1} | X_n = R).$$

This shows that $\{X_n\}$ satisfies the Markov property. \quad \Box