Problem 1 (60 points)

a. The following are statements about events $A$, $B$, $C$ with probabilities $P(A)$, $P(B)$, and $P(C)$ that are all nonzero and strictly less than 1. The events are arbitrary, unless otherwise specified.

<table>
<thead>
<tr>
<th>True/False</th>
<th>Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>✔</td>
<td>if $B \subset A$, then $P(A) \geq P(B)$</td>
</tr>
<tr>
<td>✔</td>
<td>if the probability that at most one of $A$, $B$, $C$ occurs is 1, then the events $A$, $B$, $C$ are disjoint</td>
</tr>
<tr>
<td>☐ ✔</td>
<td>if $P(A) = P(B) = P(C) = 1/3$, then all the eight regions in the Karnaugh map for $A$, $B$, $C$ have the same probability</td>
</tr>
<tr>
<td>☐ ✔</td>
<td>if $A \cap C = \emptyset$ and $P(A</td>
</tr>
<tr>
<td>☐ ✔</td>
<td>if $P(A \cup B \cup C) = P(C)$, then $P(C</td>
</tr>
<tr>
<td>☐ ✔</td>
<td>if the events $A$ and $B$ are independent and the events $A$ and $C$ are independent, then the events $B$ and $C$ are also independent</td>
</tr>
</tbody>
</table>

b. In the following statements, $X$ and $Y$ are generic random variables, unless specified otherwise in the statement. The function $F_X(u)$ is the cumulative distribution function (CDF) of $X$. The functions $f_X(u)$ and $p_X(u)$ denote the probability density function of $X$ and the probability mass function of $X$. Similarly, the functions $f_Y(u)$ and $p_Y(u)$ denote the probability density function of $Y$ and the probability mass function of $Y$.

<table>
<thead>
<tr>
<th>True/False</th>
<th>Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐ ✔</td>
<td>if $X$ is continuous, then $f_X(u) \leq 1$ for all real numbers $u$</td>
</tr>
<tr>
<td>✔ ☐</td>
<td>if $X$ is discrete and $Y = 2X$, then $p_Y(u) = p_X(u/2)$ for all $u$</td>
</tr>
<tr>
<td>✔ ☐</td>
<td>if $X$ is continuous and $f_X(u) = f_X(-u)$ for all $u$, then $F_X(0) = 0.5$</td>
</tr>
<tr>
<td>☐ ✔</td>
<td>if $X$ is discrete, $Y$ is continuous, and $p_X(\sqrt{3}) = f_Y(\sqrt{3})$ then $P(X = \sqrt{3}) = P(Y = \sqrt{3})$</td>
</tr>
</tbody>
</table>
Problem 2 (14 points)

Let $A$ be the event of rolling at least one six on six independent rolls of a fair die. Let $B$ be the event of rolling at least one double-six — getting the outcome $(6, 6)$ — on six independent rolls of a pair of fair dice. Which event is more likely? Compute both $P(A)$ and $P(B)$, then answer this question.

**Answer:** The probability of drawing a six on a single roll of a fair die is $p = \frac{1}{6}$. Therefore, the probability of not drawing any sixes on six independent rolls is $(1 - p)^6 = \left(\frac{5}{6}\right)^6$, and we have

$$P(A) = 1 - (1 - p)^6 = 1 - \frac{5^6}{6^6}$$

A very similar argument works for event $B$. The probability of drawing the outcome $(6, 6)$ on a single roll of a pair of dice is $p' = \frac{1}{36}$. Therefore, the probability of not drawing any double-sixes on six independent rolls of a pair of dice is $(1 - p')^6 = \left(\frac{35}{36}\right)^6$, and we have

$$P(B) = 1 - (1 - p')^6 = 1 - \frac{35^6}{36^6}$$

Since $\frac{35}{36} < \frac{5}{6}$, we see that $P(A) > P(B)$. This also makes sense intuitively. Rolling a six on a single roll of a die is more likely than rolling a double-six on a single roll of a pair of dice. Clearly, you need more “luck” for the latter event. Hence, this latter event is less likely to occur at least once on six independent trials.

$$\begin{align*}
P(A) &= 1 - \frac{5^6}{6^6} \\
P(B) &= 1 - \frac{35^6}{36^6} \\
\checkmark \quad A \text{ is more likely than } B \\
\square \quad B \text{ is more likely than } A
\end{align*}$$

Problem 3 (33 points)

There are $n$ students at Rumor Mill University, and Alice is one of them. On day zero, Alice hears a new rumor, of which none of the students was previously aware. On day one, Alice selects one of the other $n - 1$ students uniformly at random and tells the rumor to that student. On day two, the student who heard the rumor from Alice selects one of the other $n - 1$ students (including Alice, but not including herself/himself) at random and tells the rumor to that student. This process continues forever: on day $i$, the student who was told the rumor on day $i - 1$ selects one of the other $n - 1$ students at random and tells the rumor to that student.

For $i = 1, 2, \ldots, n$, let $A_i$ be the event that by the beginning of day $i$, there are exactly $i$ different students who have heard the rumor. Thus $P(A_1) = 1$ because by the beginning of day one, Alice is the single student who has heard the rumor. We also have $P(A_2) = 1$, because by the beginning of day two, there are necessarily two students who have heard the rumor. But $P(A_3) \neq 1$ (why?).

a. Compute $P(A_2A_1)$, $P(A_3A_2A_1)$, and $P(A_4A_3A_2A_1)$.

**Answer:** Clearly, $P(A_2A_1) = P(A_2|A_1)P(A_1) = 1$. To compute $P(A_3A_2A_1)$, proceed as follows: *given that* exactly 2 different students already heard the rumor by the beginning of day two, the probability that a new student is told the rumor on that day is $(n - 2)/(n - 1)$. That is,
of the \( n - 1 \) equally likely choices made by the student who tells the rumor on day two, exactly 
\( n - 2 \) will lead to the event \( A_3 \). Therefore, we have

\[
P(A_3A_2A_1) = P(A_3|A_2A_1)P(A_2|A_1)P(A_1) = \frac{n-2}{n-1} \cdot 1 \cdot 1 = \frac{n-2}{n-1}
\]

where we have made use of the chain rule (or multiplication rule) involving conditional probabilities (Ross, p. 63). A similar argument can be used to compute \( P(A_4A_3A_2A_1) \). Given that \( 3 \) students have heard the rumor by the beginning of day three, the probability that a new student is added to this group on day three is \( (n - 3)/(n - 1) \), because exactly \( n - 3 \) out of the \( n - 1 \) possible choices correspond to students that did not yet hear the rumor. Hence

\[
P(A_4A_3A_2A_1) = P(A_4|A_3A_2A_1)P(A_3|A_2A_1)P(A_2|A_1)P(A_1) = \frac{n-3}{n-1} \cdot \frac{n-2}{n-1}
\]

\[
P(A_2A_1) = 1 
\]

\[
P(A_3A_2A_1) = \frac{n-2}{n-1}
\]

\[
P(A_4A_3A_2A_1) = \frac{(n-2)(n-3)}{(n-1)^2}
\]

b. Compute \( P(A_iA_{i-1} \cdots A_3A_2A_1) \) for all \( i = 2, 3, \ldots, n \).

**Answer:** Computing \( P(A_iA_{i-1} \cdots A_3A_2A_1) \) for all \( i \) amounts to appropriately extending the argument used to solve part (a). Given that \( i - 1 \) students have already heard the rumor by the beginning of day \( i - 1 \), the probability that a new student is added to this group on day \( i - 1 \) is \( (n - (i-1))/(n - 1) \). Consequently, using the chain rule as before, we have

\[
P(A_iA_{i-1} \cdots A_2A_1) = P(A_i|A_{i-1} \cdots A_2A_1)P(A_{i-1}|A_{i-2} \cdots A_2A_1) \cdots P(A_2|A_1)P(A_1)
\]

\[
= \frac{n-(i-1)}{n-1} \cdot \frac{n-(i-2)}{n-1} \cdots \frac{n-3}{n-1} \cdot \frac{n-2}{n-1} = \frac{(n-2)!}{(n-i)! (n-1)^{i-2}}
\]

where we have again made use of the chain rule. Observe that for \( i = 2, 3, 4 \), the expression above produces answers that are identical to those obtained in part (a).

\[
P(A_iA_{i-1} \cdots A_3A_2A_1) = \frac{(n-2)!}{(n-i)! (n-1)^{i-2}}
\]

c. As the rumor spreads, it will eventually happen that the student who is told the rumor on a certain day has already heard the rumor. Let \( X \) be the number of the day on which this happens for the first time. Then \( X \) is a discrete random variable. Compute its probability mass function.

**Answer:** Let \( E_i \) be the event \( \{ X = i \} \). We need to find the probability of this event for all \( i \). First, suppose that the event \( A_iA_{i-1} \cdots A_3A_2A_1 \) has occurred — that is, by the beginning of day \( i \), there are exactly \( i \) different students who have heard the rumor (observe that \( A_i \subset A_{i-1} \) for all \( i \), and consequently \( A_iA_{i-1} \cdots A_3A_2A_1 = A_i \)). Then the probability that \( E_i \) will occur is \( (i-1)/(n-1) \) since of the \( i \) students who have already heard the rumor, only \( i - 1 \) can be chosen by the student who tells it (the student cannot choose herself/himself). Now, if the event \( A_iA_{i-1} \cdots A_3A_2A_1 = A_i \) did not occur, then the probability that \( E_i \) will occur is zero,
since the value of $X$ must be strictly less than $i$ in this case. Hence, using the theorem of total probability along with the result of part (b), we obtain

$$P(E_i) = P(E_i|A_i)P(A_i) + P(E_i|A_i^c)P(A_i^c) = \frac{i-1}{n-1} \cdot \frac{(n-2)!}{(n-i)! (n-1)^{i-2}}$$

To complete the solution, note that $P(E_i) = 0$ for $i = 1$ and $i > n$. The full probability mass function of $X$ is given below.

$$p_X(u) = \begin{cases} \frac{(u-1)(n-2)!}{(n-u)! (n-1)^{u-1}} & u \in \{2, 3, \ldots, n\} \\ 0 & \text{otherwise} \end{cases}$$

**Problem 4 (23 points)**

A continuous random variable $X$ is specified in terms of the following probability density function:

$$f_X(u) = \begin{cases} 2e^{-\alpha u} & \text{if } u \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

a. Determine the value of the constant $\alpha$.

**Answer:** To compute $\alpha$, we use one of the fundamental properties of probability density functions, namely

$$1 = \int_{-\infty}^{\infty} f_X(u) \, du = \int_{0}^{\infty} 2e^{-\alpha u} \, du = -\frac{2}{\alpha} e^{-\alpha u} \bigg|_{0}^{\infty} = \frac{2}{\alpha}$$

From this, it follows that $\alpha = 2$.

b. What is the cumulative distribution function (CDF) of $X$?

**Answer:** The CDF is the anti-derivative of the probability density function. That is:

$$F_X(a) = \int_{-\infty}^{a} f_X(u) \, du = \int_{0}^{a} 2e^{-2u} \, du = -e^{-2u} \bigg|_{0}^{a} = 1 - e^{-2a}$$

$$F_X(u) = \begin{cases} 1 - e^{-2u} & \text{if } u \geq 0 \\ 0 & \text{if } u < 0 \end{cases}$$

c. Compute the probability that $X > 2$, given that $X \leq 4$.

**Answer:** Using the definition of conditional probability and the result of part (b), we compute:

$$P(X > 2 \mid X \leq 4) = \frac{P(2 < X \leq 4)}{P(X \leq 4)} = \frac{F_X(4) - F_X(2)}{F_X(4)} = \frac{(1 - e^{-8}) - (1 - e^{-4})}{1 - e^{-8}}$$

$$P(X > 2 \mid X \leq 4) = \frac{e^{-4} - e^{-8}}{1 - e^{-8}}$$