1. (a) The Kolmogorov axioms are for the set function $P$:

1. $P(A) \geq 0$ for all events $A$.
2. $P(S) = 1$.
3. Suppose $A_1, A_2, ..., $ are pairwise mutually exclusive events; i.e.,

$$A_i \cap A_j = \emptyset \ \forall i \neq j.$$

Then

$$P\left( \bigcup_{i=1}^{\infty} A_i \right) = \sum_{i=1}^{\infty} P(A_i) \quad \text{“countable additivity”}.$$

(b) Mutual independence requires

$$P(A_1 \cap A_2 \cap \cdots \cap A_k) = P(A_1)P(A_2)\cdots P(A_k)$$

for any subcollection of events $A_1, A_2, ..., A_k$.

**Note:** Pairwise independence; i.e., $P(A_i \cap A_j) = P(A_i)P(A_j)$ is not strong enough. The equation above must hold for any subcollection; i.e., for any 2 events, for any 3 events, and so on. It must also hold for all $n$ events; i.e., $P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2)\cdots P(A_n)$.

2. From the LOTP, we know

$$P(A) = P(A|B)P(B) + P(A|\bar{B})P(\bar{B}).$$

Because $P(B) = 1$ by assumption, we know $P(\bar{B}) = 1 - P(B) = 0$. Therefore,

$$P(A) = P(A|B)P(B) = P(A|B)$$

because $P(B) = 1$. This means $A$ and $B$ are independent events.

(b) If $A \subset B$, then we can write

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

Note that $\bar{A} \cap B$ denotes all outcomes in $B$ but not in $A$. Furthermore $A$ and $\bar{A} \cap B$ are mutually exclusive. Therefore, by Axiom 3 (countable additivity), we have

$$P(B) = P(A) + P(\bar{A} \cap B) \implies P(\bar{A} \cap B) = P(B) - P(A).$$

(c) $A$ and $B$ are mutually exclusive means $A \cap B = \emptyset$. Using the definition of conditional probability,

$$P(A|A \cup B) = \frac{P(A \cap (A \cup B))}{P(A \cup B)}.$$

Now, $A \subset A \cup B \implies A \cap (A \cup B) = A$. Also,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B),$$

because $P(A \cap B) = P(\emptyset) = 0$ (i.e., $A$ and $B$ are mutually exclusive). Therefore,

$$P(A|A \cup B) = \frac{P(A)}{P(A) + P(B)}$$

as claimed.
3. Define the following events:

\[ A = \{ \text{patient dies} \} \]
\[ B_1 = \{ \text{patient classified critical} \} \]
\[ B_2 = \{ \text{patient classified serious} \} \]
\[ B_3 = \{ \text{patient classified stable} \} \]

We are given
\[ P(B_1) = 0.20 \]
\[ P(B_2) = 0.30 \]
\[ P(B_3) = 0.50 \]

\[ \sum \text{to 1 } (i.e., B_1, B_2, B_3 \text{ partition } S) \]

We are also given
\[ P(A|B_1) = 0.30 \]
\[ P(A|B_2) = 0.10 \]
\[ P(A|B_3) = 0.01 \]

(a) Use LOTP:

\[ P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3) \]
\[ = 0.30(0.20) + 0.10(0.30) + 0.01(0.50) = 0.095. \]

(b) Use Bayes’ Rule. We want \( P(B_1|A) \).

\[ P(B_1|A) = \frac{P(A \cap B_1)}{P(A)} = \frac{P(A|B_1)P(B_1)}{P(A)} = \frac{0.30(0.20)}{0.095} \approx 0.632. \]

4. (a) One sample point looks like this:

\[ (\begin{array}{cccc} 3 & 1 & 4 & 2 \\ E1 & E2 & E3 & E4 \end{array}) \]

This would mean that

- invitation to friend 3 went to envelope 1 (E1)
- invitation to friend 1 went to envelope 2 (E2)
- invitation to friend 4 went to envelope 3 (E3)
- invitation to friend 2 went to envelope 4 (E4).

How many sample points? This is equal to the number of ways to permute the four distinct objects “1,” “2,” “3,” and “4.” That is, there are

\[ 4! = 24 \text{ different sample points in } S. \]

The sample space consists of the 24 outcomes (sample points) like the one above.

(b) At first, one might think that \( Y \) can be 0, 1, 2, 3, or 4. Let’s calculate \( P(Y = 4) \). For \( \{Y = 4\} \) to occur, every invitation must match its envelope. This corresponds to the outcome

\[ (\begin{array}{cccc} 1 & 2 & 3 & 4 \\ E1 & E2 & E3 & E4 \end{array}) \]

There is only 1 such outcome; hence, \( P(Y = 4) = 1/24 \), assuming all 24 outcomes in \( S \) are equally likely.
We should quickly see that \( P(Y = 3) = 0 \). To see why, the event \( \{Y = 3\} \) means there were “3 matches.” If there were 3 matches, the last envelope/invitation must match too (which would be 4 matches). Therefore, \( \{Y = 3\} \) cannot occur and hence \( P(Y = 3) = 0 \).

Next, \( P(Y = 2) \). This means there are two “matches;” something like

\[
\begin{array}{cccc}
E_1 & E_2 & E_3 & E_4 \\
1 & & 3 & \\
& 2 & & \\
& & & 4
\end{array}
\]

There are \( \binom{4}{2} = 6 \) ways to choose 2 envelopes whose invitations match. Once these 2 envelopes/invitations matches are selected, the remaining invitations must be put in the wrong envelope. Therefore, there are 6 ways the event \( \{Y = 2\} \) can occur and thus \( P(Y = 2) = 6/24 \), assuming all 24 outcomes in \( S \) are equally likely.

Next, \( P(Y = 1) \). Exactly 1 match. Use the basic rule of counting:

\[
\begin{align*}
 n_1 &= \text{number of ways to select 1 envelope/invitation match} = 4 \\
 n_2 &= \text{number of ways 2nd invitation can be assigned to wrong envelope} = 2.
\end{align*}
\]

The third and fourth invitations must be assigned to the wrong envelope as well. Therefore, the number of ways \( \{Y = 1\} \) can occur is \( n_1 \times n_2 = 8 \). Therefore, \( P(Y = 1) = 8/24 \), assuming all 24 outcomes in \( S \) are equally likely.

Finally, we can get \( P(Y = 0) \) by subtraction. We know

\[
P(Y = 0) + P(Y = 1) + P(Y = 2) + P(Y = 4) = 1.
\]

Therefore,

\[
P(Y = 0) = 1 - \frac{8}{24} - \frac{6}{24} - \frac{1}{24} = \frac{9}{24}.
\]

The pmf of \( Y \) can be written as

<table>
<thead>
<tr>
<th>( y )</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_Y(y) )</td>
<td>9/24</td>
<td>8/24</td>
<td>6/24</td>
<td>1/24</td>
</tr>
</tbody>
</table>

5. (a) \( Y \) is a discrete random variable with four possible values \( (0, 2, 4, \text{ and } 10) \). We know

\[
p_Y(0) + p_Y(2) + p_Y(4) + p_Y(10) = 1 \implies 0.8 + 2c + 4c + 10c = 1 \implies 16c = 0.2 \implies c = \frac{1}{80}.
\]

The pmf of \( Y \) can be written as

<table>
<thead>
<tr>
<th>( y )</th>
<th>0</th>
<th>2</th>
<th>4</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_Y(y) )</td>
<td>0.8</td>
<td>2/80</td>
<td>4/80</td>
<td>10/80</td>
</tr>
</tbody>
</table>

The expected loss is

\[
E(Y) = 0(0.8) + 2 \left( \frac{2}{80} \right) + 4 \left( \frac{4}{80} \right) + 10 \left( \frac{10}{80} \right) = \frac{120}{80} = 1.5.
\]
(b) Note the support is \( R = \{0, 2, 4, 10\} \). The mgf of \( Y \) is

\[
m_Y(t) = E(e^{ty}) = \sum_{y \in R} e^{ty} p_Y(y)
\]
\[
= e^{t(0)} \left( 0.8 \right) + e^{t(2)} \left( \frac{2}{80} \right) + e^{t(4)} \left( \frac{4}{80} \right) + e^{t(10)} \left( \frac{10}{80} \right)
\]
\[
= 0.8 + \frac{2}{80} e^{2t} + \frac{4}{80} e^{4t} + \frac{10}{80} e^{10t}.
\]

6. (a) The mgf of \( Y \) is

\[
m_Y(t) = E(e^{ty}) = \sum_{y=0}^{\infty} e^{ty} p_Y(y) = \sum_{y=0}^{\infty} e^{ty} \left( \frac{1}{4} \right)^y = \frac{3}{4} \sum_{y=0}^{\infty} \left( \frac{e^t}{4} \right)^y.
\]

The last sum is an infinite geometric sum with common ratio \( r = e^t/4 \). Provided that \( r < 1 \); i.e.,
\[
e^t/4 < 1 \iff t < \ln 4
\]

this sum converges and hence \( m_Y(t) \) exists. We have

\[
m_Y(t) = \frac{3}{4} \left( \frac{1}{1 - e^t/4} \right) = \frac{3}{4} \left( \frac{4}{4 - e^t} \right) = \frac{3}{4 - e^t}.
\]

(b) To take derivatives, write \( m_Y(t) = 3(4 - e^t)^{-1} \). The first derivative is

\[
\frac{d}{dt} m_Y(t) = 3(-1)(4 - e^t)^{-2}(-e^t) = 3e^t(4 - e^t)^{-2}.
\]

Thus,

\[
E(Y) = \left. \frac{d}{dt} m_Y(t) \right|_{t=0} = 3e^0(4 - e^0)^{-2} = \frac{1}{3}.
\]

The second derivative is

\[
\frac{d^2}{dt^2} m_Y(t) = \frac{d}{dt} \left( 3e^t(4 - e^t)^{-2} \right) = 3e^t(4 - e^t)^{-2} + 3e^t(-2)(4 - e^t)^{-3}(-e^t)
\]
\[
= \frac{3e^t}{(4 - e^t)^2} + \frac{6e^{2t}}{(4 - e^t)^3}.
\]

Thus,

\[
E(Y^2) = \left. \frac{d^2}{dt^2} m_Y(t) \right|_{t=0} = \frac{3e^0}{(4 - e^0)^2} + \frac{6e^{2(0)}}{(4 - e^0)^3} = \frac{3}{4} + \frac{6}{27} = \frac{15}{27}.
\]

Finally, from the variance computing formula,

\[
V(Y) = E(Y^2) - [E(Y)]^2 = \frac{15}{27} - \left( \frac{1}{3} \right)^2 = \frac{4}{9}.
\]