1. Probability / Stats

(a) Assume the following data were drawn independently at random from a discrete probability distribution. What are the sample frequencies?

\[
\begin{array}{ccccccccccc}
D & B & B & C & B & C & C & B & C & B \\
C & C & B & B & D & C & D & A & C & B \\
\end{array}
\]

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td></td>
</tr>
</tbody>
</table>

(b) The tables give below give symbol probabilities for two hypothesized distributions that could explain the data above. What is the likelihood that the data was drawn from \( H_0 \)? From \( H_1 \)?

\[
\begin{array}{|c|c|}
\hline
\text{Symbol} & \text{Probability} \\
\hline
A & 0.1 \\
B & 0.4 \\
C & 0.3 \\
D & 0.2 \\
\hline
\end{array}
\quad
\begin{array}{|c|c|}
\hline
\text{Symbol} & \text{Probability} \\
\hline
A & 0.25 \\
B & 0.25 \\
C & 0.25 \\
D & 0.25 \\
\hline
\end{array}
\]

Parameters for \( H_1 \)

Parameters for \( H_0 \)

(c) Show how you could calculate the KL divergence from the distribution for \( H_0 \) to the observed distribution of symbols in part (a).