Exercise 16
Let \((\Omega, P)\) be a probability space.

1. Show that the law of total probability also holds for conditional probabilities, i.e. that
\[
P(A \mid C) = \sum_{i \in I} P(A, B_i \mid C)
\]
for events \(A\) and \(C\), and a countable family \((B_i \mid i \in I)\) which partitions the sample space.

2. Show that
\[
P(A, B \mid C) = P(B \mid C) \cdot P(A \mid B, C)
\]
for events \(A, B,\) and \(C\). □

Exercise 17
Let \(m, n \in \mathbb{N}\), and \(f: \{1, \ldots, m\} \times \{1, \ldots, n\} \rightarrow \mathbb{R}\) be a mapping. Show the following equality.
\[
\sum_{a \in \{1, \ldots, n\}} \prod_{i \in \{1, \ldots, m\}} f(i, a(i)) = \prod_{i \in \{1, \ldots, m\}} \sum_{j \in \{1, \ldots, n\}} f(i, j)
\]
Apply this equation to IBM model 1.

Note: For two sets \(A\) and \(B\) the notion \(B^A\) is defined as the set of all mappings from \(A\) to \(B\), i.e. \(\{f \mid f: A \rightarrow B\}\). □

Exercise 18 (IBM model 1)
Consider the following dictionary.

\[
\begin{array}{c|cccc}
  t(f \mid e) & kra & ban & las & gha \\
  \hline
da & 0.2 & 0.4 & 0.4 & 0 \\
\varepsilon & 0 & 0.1 & 0.8 & 0.1 \\
ur & 0.3 & 0.4 & 0.25 & 0.25 \\
fur & 0.4 & 0.3 & 0.1 & 0.2 \\
\end{array}
\]

1. Additionally, consider the following length model: \(\varepsilon(m \mid l) = 0.5\) if \(m = l\), \(\varepsilon(m \mid l) = 0.25\) if \(|m - l| = 1\), and \(\varepsilon(m \mid l) = 0\) otherwise. Determine \(P(f \mid e)\) for the following sentence pairs by means of IBM model 1:
• \( f = \text{kra las gha}, \ e = \text{du su ur}, \)
• \( f = \text{kra las gha}, \ e = \text{du su}, \)
• \( f = \text{gha gha}, \ e = \text{su du su}. \)

2. Suppose that \( \varepsilon(m \mid l) = 2^{-m} (m \geq 1). \) Determine \( \arg\max_f P(f \mid e) \) for each of the following sentences:
- \( e = \text{du}, \)
- \( e = \text{du su ur}, \)
- \( e = \text{fur du fur su fur}. \)

Now assume that \( \varepsilon(m \mid l) = 1 \) if \( l = m \) and \( \varepsilon(m \mid l) = 0 \) otherwise. Try to determine the best translations for the above sentences.

Exercise 19 (Bigrams)
Let \( E = \{\text{du}, \text{su}, \text{ur}, \text{fur}, \text{mu}\} \) and \( F = \{\text{kra, ban, las, gha, ra}\}. \) Consider the following dictionary:

\[
\begin{array}{l|lllll}
 & \text{kra} & \text{ban} & \text{las} & \text{gha} & \text{ra} \\
\hline
\text{du} & 1 & 0 & 0 & 0 & 0 \\
\text{su} & 0 & 1 & 0 & 0 & 0 \\
\text{e ur} & 0 & 0 & 1 & 0 & 0 \\
\text{fur} & 0 & 0 & 0 & 1 & 0 \\
\text{mu} & 0 & 0 & 0 & 0 & 1 \\
\end{array}
\]

Let \( \varepsilon(m \mid l) = 1 \) if \( l = m \) and \( \varepsilon(m \mid l) = 0 \) otherwise. Decode the sentence “kra ban las gha ra” using the following bigram model:

\[
\begin{array}{l|llllll}
 & \text{du} & \text{su} & \text{ur} & \text{fur} & \text{mu} & \#
\hline
\text{du} & 0 & 0.2 & 0.1 & 0.3 & 0.2 & 0.2 \\
\text{su} & 0.3 & 0 & 0 & 0.3 & 0.3 & 0.1 \\
\text{e ur} & 0.1 & 0 & 0 & 0.7 & 0 & 0.2 \\
\text{fur} & 0.3 & 0.2 & 0.2 & 0 & 0 & 0.3 \\
\text{mu} & 0.2 & 0.5 & 0 & 0 & 0 & 0.3 \\
\# & 1 & 0 & 0 & 0 & 0 & 0 \\
\end{array}
\]