Task 2.1 (Entropy of the geometric distribution)
For the analysis of data transmission systems, it is usually assumed that the length of the transmitted messages are geometrically distributed. If $N$ is the number of bits in a message, then it holds that

$$f_N(n) = p(1-p)^{n-1} \text{ for } n = 1, 2, \ldots \text{ and } 0 < p < 1.$$ 

a) Show that the expected value of the random variable $N$ is $\mu_N = 1/p$.

b) Calculate the uncertainty $H(N)$ about the length $N$ of the message subject to $p$ and simplify this expression as much as possible using the expected value $\mu_N$ and the binary entropy function

$$h(p) = -p \log_2(p) - (1-p) \log_2(1-p).$$

Task 2.2 (Conditional entropies)
Let the joint distribution $f_{XY}(x, y)$ of two random variables $X$ and $Y$ be given by

$$
\begin{array}{c|cc}
X \backslash Y & 0 & 1 \\
\hline
0 & \frac{1}{3} & \frac{1}{3} \\
1 & 0 & \frac{1}{3}
\end{array}
$$

Find:

a) $H(X), H(Y)$

b) $H(X|Y), H(Y|X)$

c) $H(X,Y)$

d) $H(Y) - H(Y|X)$

Task 2.3 (Conditional mutual information)
Give examples of joint random variables $X, Y,$ and $Z$ such that

a) $I(X;Y|Z) < I(X;Y)$.

b) $I(X;Y|Z) > I(X;Y)$.
Task 2.4 (Entropy and mutual information)
Let $\Omega = \{(000,),(011),(101),(110)\}$ be a set, $(X_1X_2X_3)$ a vectorial-valued random variable with $P((X_1, X_2X_3) = \omega) = \frac{1}{4}$ $\forall \omega \in \Omega$. Determine:

(i) $H(X_1)$

(ii) $H(X_1X_2)$

(iii) $H(X_2|X_1)$

(iv) $H(X_1X_2X_3)$

(v) $H(X_3|X_1X_2)$

(vi) $H(X_3)$

(vii) $I(X_1;X_3)$

(viii) $I(X_1X_2;X_3)$

Task 2.5 (Kullback-Leibler distance and average mutual information)
Let $X$ and $Y$ be two random variables with the same set of events. The Kullback-Leibler distance is defined as

$$D_{KL}(f_X(x)||f_Y(x)) = \sum_{x \in supp f_X(x)} f_X(x) \log \frac{f_X(x)}{f_Y(x)} = E\left[\log \frac{f_X(x)}{f_Y(x)}\right].$$

Let $f_{XY}(x,y)$ be the joint distribution and $f_X(x)$ and $f_Y(y)$ be the marginal distributions. Show that the average mutual information of $X$ and $Y$ is given by

$$I(X;Y) = D_{KL}(f_{XY}(x,y)||f_X(x)f_Y(y)).$$