Problem 2.1(a)
The pmf is \( P_X(x) = (1/2)^x \) for \( x \geq 1 \). There, by definition, the entropy is

\[
H(X) = - \sum_{x=1}^{\infty} P_X(x) \log P_X(x) \tag{1}
\]

\[
= - \sum_{x=1}^{\infty} \left( \frac{1}{2} \right)^x \log \left( \frac{1}{2} \right)^x \tag{2}
\]

\[
= \sum_{x=1}^{\infty} x \left( \frac{1}{2} \right)^x \log 2 \tag{3}
\]

\[
= 2 \text{ (bits)} \tag{4}
\]

Problem 2.2
(a) Since it is a one-to-one mapping between \( X \) and \( Y \), \( H(X) = H(Y) \).

(b) Here \( y \) is uniquely determined by the value of \( x \), which means \( H(Y|X) = 0 \). Then, in general, we have

\[
H(Y) = H(X) + H(Y|X) - H(X|Y) \tag{5}
\]

\[
= H(X) - H(X|Y) \tag{6}
\]

\[
\leq H(X) \tag{7}
\]

where the inequality is due to \( H(X|Y) \geq 0 \).

Comment: Under certain conditions, the equality holds. For example, \( X = 0 \) (constant). Then, \( Y = \cos 0 = 1 \) with probability 1, and \( H(X) = H(Y) = 0 \).

Problem 2.3
It is clear that \( H(p_1, p_2, \cdots, p_n) \geq 0 \). So 0 is always a lower bound. To achieve it, we must have

\[
- \sum_{i=1}^{n} p_i \log p_i = 0 \tag{8}
\]

which implies \( p_i = 0 \) or 1 for all \( i \). We also have the constraint \( \sum_{i=1}^{n} p_i = 1 \). So the distributions which yield the minimum 0 are

\[
(1, 0, 0, \cdots, 0), (0, 1, 0, \cdots, 0), \cdots, (0, 0, 0, \cdots, 1) \tag{9}
\]

Problem 2.6
(a) Let \( X \) be a random variable with \( Ber(1/2) \), i.e., \( P(X = 0) = P(X = 1) = 1/2 \). We also choose \( Y = Z = X \). Then, we have

\[
I(X;Y|Z) = H(X|Z) - H(X|Y, Z) = 0 \tag{10}
\]

\[
I(X;Y) = H(X) - H(X|Y) = 1 > I(X;Y|Z) \tag{11}
\]
(b) Let $X$ and $Y$ be two independent random variables with $Ber(1/2)$, and $Z = X + Y$. Then, it is easy to check that

$$ I(X;Y|Z) = H(X|Z) - H(X|Y,Z) = H(X|Z) > 0 = I(X;Y) \quad (12) $$

**Problem 2.9**

(a) We need to check all the properties in the definition:

- $\rho(X,Y) = H(X|Y) + H(Y|X) \geq 0 + 0 = 0$.
- $\rho(X,Y) = H(X|Y) + H(Y|X) = H(Y|X) + H(X|Y) = \rho(Y,X)$.

$$ \rho(X,Y) + \rho(Y,Z) = H(X|Y) + H(Y|Z) + H(Z|Y) \quad (13) $$

$$ = H(X|Y) + H(Y|X) + H(Y|Z) + H(Z|Y) \quad (14) $$

$$ \geq H(X|Y,Z) + H(Y|Z) + H(Z|Y,X) + H(Y|X) \quad (15) $$

$$ = H(X,Y,Z) + H(Y|X) \quad (16) $$

$$ \geq H(X|Z) + H(Z|X) \quad (17) $$

$$ = \rho(X,Z) \quad (18) $$

- If there is a one-to-one mapping between $X,Y$, then $H(X|Y) = H(Y|X) = 0$, i.e., $\rho(X,Y) = 0$. On the other hand, if $\rho(X,Y) = 0$, then we must have $H(X|Y) = H(Y|X) = 0$. By definition,

$$ 0 = H(X|Y) = - \sum_{y \in Y} P_Y(y) \sum_{x \in X} P_{X|Y}(x|y) \log P_{X|Y}(x|y) \quad (19) $$

which means that, for each $y \in Y$ with $P_Y(y) > 0$, we must have either $P_{X|Y}(x|y) = 0$ or $P_{X|Y}(x|y) = 1$ for all $x$. Combined with $\sum_{x \in X} P_{X|Y}(x|y) = 1$, we can conclude that for each $y (P_Y(y) > 0)$ there exists one and only one $x_y$ such that $p(x_y|y) > 0$. The same argument can be applied to $H(Y|X) = 0$, which indicates that there is a one-to-one mapping between $X,Y$, i.e., $X \equiv Y$.

(b)

$$ \rho(X,Y) = H(X|Y) + H(Y|X) \quad (20) $$

$$ = H(X,Y) - H(Y) + H(X,Y) - H(X) \quad (21) $$

$$ = 2H(X,Y) - H(Y) - H(X) \quad (22) $$

$$ = H(X,Y) + [H(X,Y) - H(X) - H(Y)] \quad (23) $$

$$ = H(X,Y) - I(X;Y) \quad (24) $$

$$ = H(X) + H(Y) + [H(X,Y) - H(X) - H(Y)] - I(X;Y) \quad (25) $$

$$ = H(X) + H(Y) - 2I(X;Y) \quad (26) $$

**Problem 2.13**

Let $f(x) = \ln x - 1 + \frac{1}{x}$. For $x > 0$, the derivative of $f$ exists and is

$$ f'(x) = \frac{1}{x} - \frac{1}{x^2} \quad (28) $$

which implies that for $x > 0$, $x = 1$ is a global minimum or maximum point. Check the value of the second-order derivative at $x = 1$

$$ f''(x) \bigg|_{x=1} = \left(- \frac{1}{x^2} + 2 \frac{1}{x^3}\right) \bigg|_{x=1} = 1 > 0 \quad (29) $$
which means that $x = 1$ is a global minimum and $f(x) \geq f(1) = 0$ for $x > 0$, i.e.,

$$
\ln x \geq 1 - \frac{1}{x}, \quad (x > 0)
$$

(30)

**Problem 2.25**

(a) Check Problem 2.6 for the example of $I(X; Y; Z) < 0$.

(b) Starting from Eqn.(34)

$$
I(X; Y; Z) = H(X, Y, Z) + I(X; Y) - H(X|Z) - H(Y, Z)
$$

(34)

$$
= H(X, Y, Z) + I(X; Y) - H(X, Z) + H(Z) - H(Y, Z)
$$

(38)

$$
= H(X, Y, Z) - H(X, Y) - H(Y, Z) - H(X, Z) + H(Z) + H(X, Y) + I(X; Y)
$$

(39)

$$
= H(X, Y, Z) - H(X, Y) - H(Y, Z) - H(X, Z) + H(Z) + H(X) + H(Y)
$$

(40)

**Prove identities**

(a)

$$
I(X, Y; Z) = H(X, Y) - H(X|Z)
$$

(41)

$$
= H(X) + H(Y|X) - H(X|Z) - H(Y|X, Z)
$$

(42)

$$
= H(X) - H(X|Z) + H(Y|X) - H(Y|X, Z)
$$

(43)

$$
= I(X; Z) + I(Y; Z|X)
$$

(44)

(b)

$$
I(X, Y; Z|U) = H(X, Y|U) - H(X, Y|Z, U)
$$

(45)

$$
= H(X|U) + H(Y|X, U) - H(X|Z, U) - H(Y|X, Z, U)
$$

(46)

$$
= H(X|U) - H(X|Z, U) + H(Y|X, U) - H(Y|X, Z, U)
$$

(47)

$$
= I(X; Z|U) + I(Y; Z|X, U)
$$

(48)