Problem 1: Suppose P is a probability measure defined on $S = \{1, 2, 3, 4, 5\}$ with mass function

s	$p(s)$
1	0.1
2	0.3
3	0.2
4	0.3
5	0.1

Compute the information content $I(s)$ of each sample point and the entropy $H(P)$.

Problem 2: Let $X \sim \mathcal{B}er(\theta)$ for $\theta \in [0, 1]$. Compute a formula for $H(X)$ in terms of θ.

Problem 3: Suppose P and Q are probability measures defined on $S = \{1, 2, 3, 4, 5\}$ with mass functions

s	$p(s)$	$q(s)$
1	0.1	0.05
2	0.3	0.15
3	0.2	0.7
4	0.3	0.03
5	0.1	0.07

Compute the cross entropies $H_P(Q)$ and $H_Q(P)$.

Problem 4: Suppose P and Q are probability measures defined on $S = \{1, 2, 3, 4, 5\}$ with mass functions

s	$p(s)$	$q(s)$
1	0.1	0.05
2	0.3	0.15
3	0.2	0.7
4	0.3	0.03
5	0.1	0.07

Compute $D(P \parallel Q)$ and $D(Q \parallel P)$.

Problem 5: While it is **not** true that $D(P \parallel Q) = D(Q \parallel P)$ for **all** P and Q, there are examples of special distributions P and Q for which equality does hold. Find examples.

Problem 6: For each $\phi \in [0,1]$, consider the communication channel with transition matrix

$$\mathbf{K} = \begin{bmatrix} 1-\phi & \phi \\ \phi & 1-\phi \end{bmatrix}$$

These are called *binary symmetric channels*. Suppose that $X \sim \mathcal{B}er(\alpha)$ for some $\alpha \in [0,1]$, with the range of X enumerated as $x_0 = 0$ and $x_1 = 1$. Show that X and the communication channel determine a random variable Y with range $y_0 = 0$ and $y_1 = 1$. Determine its distribution.

Problem 7: Suppose X and Y are Bernoulli random variables with joint mass function given by

$p(x,y)$	$y=0$	$y=1$
$x=0$	0.3	0.1
$x=1$	0.36	0.24

(a) Compute the transition matrix of the communication channel induced from X and Y.

(b) Compute the mutual information $I(X,Y)$.