3. **[7 Points]** Let X and Y be random variables taking the values $\{1, \cdots, n\}$. Recall the definition of mutual information between X and Y:

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

Using the conclusion of part 1), show that the mutual information is always non-negative. Show all steps of your work.

$$H(Y) = -\sum_{j=1}^{n} P(y) \log P(y)$$

$$= -\sum_{j=1}^{n} \sum_{i=1}^{n} P(x,y) \log P(y)$$

$$H(Y|X) = \sum_{i=1}^{n} P(X=i) H(Y|X=i)$$

$$= \sum_{i=1}^{n} P(X=i) \times \left(-\sum_{j=1}^{n} P(Y=j|X=i) \log_2 P(Y=j|X=i)\right)$$

$$= -\sum_{i=1}^{n} \sum_{j=1}^{n} P(X=i)P(Y=j|X=i) \log_2 P(Y=j|X=i)$$

$$= -\sum_{i=1}^{n} \sum_{j=1}^{n} P(X=i, Y=j) \log_2 P(Y=j|X=i)$$

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

$$= -\sum_{i=1}^{n} \sum_{j=1}^{n} P(X=i,Y=j) \log_2 P(Y=j) + \sum_{i=1}^{n} \sum_{j=1}^{n} P(X=i,Y=j) \log_2 P(Y=j|X=i)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} P(X=i,Y=j) \left[\log_2 P(Y=j|X=i) - \log_2 P(Y=j)\right]$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} P(X=i,Y=j) \log_2 \frac{P(Y=j|X=i)}{P(Y=j)}$$