Joint-conditional entropy and mutual information

October 2, 2023

1 Entropy, joint entropy, conditional entropy, and mutual information

Consider the discrete source (the transmitter) X with $n = 3$ symbols with $p(x_i) = [0.125, 0.25, 0.625]$. Compute the entropy of the source H_X

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log_2(p(x_i))$$

and the maximal entropy in the case of a uniform distribution.

```
Entropy of the source H_X = 1.29879 bit
Max. entropy = 1.58496 bit
```

Consider a receiver Y with $m = 4$ symbols with the following conditional probabilities

$$\begin{gathered}
p(y_1|x_1) = 0.75 \\
p(y_2|x_1) = 0.25 \\
p(y_2|x_2) = 0.5 \\
p(y_3|x_2) = 0.5 \\
p(y_3|x_3) = 0.25 \\
p(y_4|x_3) = 0.75
\end{gathered}$$

The other probabilies are equal to zero. Note that for each transmitted symbol x_i, it has been assured that

$$\sum_{j=1}^{m} p(y_j|x_i) = 1 \qquad \forall x_i \text{ with } i \in [1, n]$$

as the sum over all possible probabilities must be equal to one.

```
P_Y_cond_X:
[[0.75 0.25 0.   0.  ]
 [0.   0.5  0.5  0.  ]
 [0.   0.   0.25 0.75]]
```

Check that row of the matrix represents a probability.

```
[1. 1. 1.]
```

Compute the joint probabilities using Bayes' rule

$$p(x_i, y_j) = p(y_j|x_i)p(x_i),$$

the probability of y_j as marginal of $p(x_i, y_j)$

$$p(y_j) = \sum_{i=1}^{n} p(x_i, y_j)$$

and the conditional probabilities $p(x_i|y_j)$ using Bayes' rule

$$p(x_i|y_j) = \frac{p(x_i, y_j)}{p(y_j)} = \frac{p(y_j|x_i)p(x_i)}{p(y_j)}.$$

```
P_joint:
[[0.09375 0.03125 0.      0.     ]
 [0.      0.125   0.125   0.     ]
 [0.      0.      0.15625 0.46875]]
P_y:
[0.09375 0.15625 0.28125 0.46875]
P_X_cond_Y:
[[1.         0.2        0.         0.        ]
 [0.         0.8        0.44444444 0.        ]
 [0.         0.         0.55555556 1.        ]]
```

Compute the entropy $H(X)$

$$H(Y) = -\sum_{j=1}^{m} p(y_j) \log_2(p(y_j))$$

and the joint entropy

$$H(X, Y) = -\sum_{i=1}^{n} \sum_{j=1}^{m} p(x_i, y_j) \log p(x_i, y_j)$$

and the the conditional entropies

$$H(Y|X) = -\sum_{i=1}^{n} \sum_{j=1}^{m} p(x_i, y_j) \log p(y_j|x_i)$$

and

$$H(X|Y) = -\sum_{i=1}^{n} \sum_{j=1}^{m} p(x_i, y_j) \log p(x_i|y_j).$$

```
Entropy H_Y = 1.76571 bit
Joint entropy: H(X,Y) = 2.15725 bit
Conditional entropy: H(Y|X) = 0.85846 bit
Conditional entropy: H(X|Y) = 0.39154 bit
```

Compute the mutual information between X and Y

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

```
Mutual information: I(X;Y) = 0.90725 bit
```

Verify the following theoretical results using the numerical results and the venn-diagrams

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

$$H(Y|X) \leq H(Y)$$

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

$$H(X|Y) \leq H(X)$$

$$0 \leq \max\left[H(X), H(Y)\right] \leq H(X,Y) \leq H(X) + H(Y)$$

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

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

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

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

$$0 \leq I(X;Y) \leq H(X)$$

```
H_X = 1.29879
H_Y = 1.76571
H(X,Y) = 2.15725
H(Y|X) = 0.85846
H(X|Y) = 0.39154
I(X;Y) = 0.90725
[H_Y_cond_X, H_joint-H_X]:
[0.8584585930558105, 0.8584585924787329]
[H_X_cond_Y, H_joint-H_Y]:
[0.3915414064482629, 0.3915414057900335]
[I_X_Y, H_X-H_X_cond_Y, H_X+H_Y-H_joint, H_Y-H_Y_cond_X, H_joint-H_X_cond_Y-
H_Y_cond_X]:
[0.9072535342471355, 0.9072535342471355, 0.9072535349053652, 0.9072535343282874,
0.9072535336700579]
```