ECS 452: In-Class Exercise # 5 Sol

Instructions

1. Separate into groups of no more than three students each. The group cannot be the same as any of your former groups. Only one submission is needed for each group.
2. Write down all the steps that you have done to obtain your answers. You may not get full credit even when your answer is correct without showing how you get your answer.
3. Do not panic.

1. In each part below, we consider a random variable $X$ which has five possible values. The probability for each possible value is listed in the provided table. Calculate the corresponding entropy value.

a.

<table>
<thead>
<tr>
<th>$x$</th>
<th>a</th>
<th>e</th>
<th>$\ell$</th>
<th>n</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p(x)$</td>
<td>0.25</td>
<td>0.25</td>
<td>0.125</td>
<td>0.25</td>
<td>0.125</td>
</tr>
</tbody>
</table>

$H(X) = -\sum_x p(x) \log_2 p(x) = -(3 \times 0.25 \times \log_2 0.25 + 2 \times 0.125 \times \log_2 0.125)$

$= -(3 \times \frac{1}{4} \times \log_2 \frac{1}{4} + 2 \times \frac{1}{8} \times \log_2 \frac{1}{8}) = 3 \times \frac{1}{4} \times 2 + 2 \times \frac{1}{8} \times 3 = \frac{6}{4} + \frac{3}{4} = \frac{9}{4} = 2.25 \text{ [bits]}$

b.

<table>
<thead>
<tr>
<th>$x$</th>
<th>a</th>
<th>e</th>
<th>$\ell$</th>
<th>n</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p(x)$</td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.3</td>
</tr>
</tbody>
</table>

$H(X) = -\sum_x p(x) \log_2 p(x) = -(0.1 \times \log_2 0.1 + 3 \times 0.2 \times \log_2 0.2 + 0.3 \times \log_2 0.3)$

$\approx 0.1 \times 3.3219 + 3 \times 0.2 \times 2.3219 + 0.3 \times 1.7370 \approx 0.3322 + 3 \times 0.4644 + 0.5211 \approx 2.2464 \text{ [bits]}$

2. [ENRPr] In each row of the following table, compare the quantity in the first column with the one in the third column by writing “>”, “=” or “<” in the second column. Watch out for approximation error.

<table>
<thead>
<tr>
<th></th>
<th>$H(p)$</th>
<th>$p$</th>
<th>$H(p)$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>$H(p) \approx 0.9710$ when $p = [0.4,0.6]$</td>
<td>$&gt;$</td>
<td>$H(p) \approx 0.8813$ when $p = [0.3,0.7]$</td>
<td></td>
</tr>
<tr>
<td>(b)</td>
<td>$H(p) \approx 0.9710$ when $p = [0.4,0.6]$</td>
<td>$&lt;$</td>
<td>$H(p) \approx 1.2955$ when $p = [0.1,0.3,0.6]$</td>
<td></td>
</tr>
<tr>
<td>(c)</td>
<td>$H(p) \approx 1.8464$ when $p = [0.1,0.2,0.3,0.4]$</td>
<td>$= H(p) \approx 1.8464$ when $p = [0.1,0.3,0.2,0.4]$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

First, note that the arguments of the entropy function are vectors. The vectors themselves represent probability mass functions. (Note that the sum of the elements in each vector is one.) Given a probability vector $p$,

$H(p) \equiv -\sum p_i \log_2 p_i$.

We can try to calculate the entropy values directly to compare them. This is done in the table above. Alternatively, we can avoid directly computing the values by the following observation:

a. When $p = [p_1,p_2]$, we have

$H(p) \equiv -p_1 \log_2 p_1 - p_2 \log_2 p_2 = -p_1 \log_2 p_1 + (1-p_1) \log_2 (1-p_1)$.

This is the same as the binary entropy function $H(p)$ evaluated at $p = p_1$. In class, we have seen the plot of $H(p)$ as a function of $p$. When the value of $p$ is closer to 0.5, the entropy value is larger. Here, 0.4 is closer to 0.5 than 0.3 is.

b. We are comparing $-0.4 \log_2 0.4 = -0.6 \log_2 0.6$ and $-0.1 \log_2 0.1 - 0.3 \log_2 0.3 - 0.6 \log_2 0.6$.

Both quantities have the term $-0.6 \log_2 0.6$; so, we can ignore this term.

Now, note that 0.4 = 0.1 + 0.3. So,

$-0.4 \log_2 0.4 = -(0.1 + 0.3) \log_2 0.4 = -0.1 \log_2 0.4 - 0.3 \log_2 0.4$.

Because $\log_2 (x)$ is an increasing function, we know that $\log_2 0.4$ is greater than both $\log_2 0.1$ and $\log_2 0.3$.

With the negative signs, $-\log_2 0.4$ is less than both $-\log_2 0.1$ and $-\log_2 0.3$.

Therefore, continued from $0$, we have

$-0.4 \log_2 0.4 = -0.1 \log_2 0.4 - 0.3 \log_2 0.4 < -0.1 \log_2 0.1 - 0.3 \log_2 0.3$.

c. The elements in vectors $p$ are the same on both sides except that the ordering is different. By the commutative property of addition, the resulting entropy values should be the same.