CS 260: Assignment #1

Due on Thursday, Jaunary 14, 2016

Prof. Wim van Dam

Chad Spensky

1 Question 1

The problem asks to prove that:

$$H(X) \leq 2log(\sum_{x \in \mathcal{X}} \sqrt{p(x)}$$

This can be done in the following way:

$$\begin{aligned}
H(X) &= \sum_{x \in \mathcal{X}} -p(x)log(p(x)) && (1)\\
&= \sum_{x \in \mathcal{X}} -2\frac{1}{2}p(x)log(p(x)) && (2)\\
&= \sum_{x \in \mathcal{X}} 2p(x)log(\frac{1}{\sqrt{p(x)}}) && (3)\\
&= E_p[2log(\frac{1}{\sqrt{p(x)}}))] && (4)\\
&\leq 2log[E_p(\frac{1}{\sqrt{p(x)}})] && (5)\\
2log[E_p(\frac{1}{\sqrt{p(x)}})] &= 2log[\sum_{x \in \mathcal{X}} \frac{p(x)}{\sqrt{p(x)}}] && (6)\\
&= 2log[\sum_{x \in \mathcal{X}} \sqrt{p(x)}] && (7)
\end{aligned}$$

where 5 is an applications of Jensen's inequality. $\square$

2 Question 2

In this question we are asked for the various entropies give all of the joint probabilities.

$$Pr[\mathcal{X} = 0, \mathcal{Y} = a] = 0.15 \quad Pr[\mathcal{X} = 0, \mathcal{Y} = b] = 0.3 \quad Pr[\mathcal{X} = 0, \mathcal{Y} = c] = 0.05$$
$$Pr[\mathcal{X} = 1, \mathcal{Y} = a] = 0.25 \quad Pr[\mathcal{X} = 1, \mathcal{Y} = b] = 0.15 \quad Pr[\mathcal{X} = 1, \mathcal{Y} = c] = 0.1$$

First, we'll define the individual probability distributions:

$$\begin{aligned}
Pr[\mathcal{X} = 0] = Pr[\mathcal{X} = 0, \mathcal{Y} = a] + Pr[\mathcal{X} = 0, \mathcal{Y} = b] + Pr[\mathcal{X} = 0, \mathcal{Y} = c] \quad &= 0.5 \\
Pr[\mathcal{X} = 1] = Pr[\mathcal{X} = 1, \mathcal{Y} = a] + Pr[\mathcal{X} = 1, \mathcal{Y} = b] + Pr[\mathcal{X} = 1, \mathcal{Y} = c] \quad &= 0.5
\end{aligned}$$

$$\begin{aligned}
Pr[\mathcal{Y} = a] = Pr[\mathcal{X} = 0, \mathcal{Y} = a] + Pr[\mathcal{X} = 1, \mathcal{Y} = a] \quad &= 0.4 \\
Pr[\mathcal{Y} = b] = Pr[\mathcal{X} = 0, \mathcal{Y} = b] + Pr[\mathcal{X} = 1, \mathcal{Y} = b] \quad &= 0.45 \\
Pr[\mathcal{Y} = c] = Pr[\mathcal{X} = 0, \mathcal{Y} = c] + Pr[\mathcal{X} = 1, \mathcal{Y} = c] \quad &= 0.15
\end{aligned}$$

First, we will calculate the basic entropies:

$$\begin{aligned}
H(X) &= \sum_{x \in \mathcal{X}} -p(x) log(p(x)) \\
&= -0.5 * log(0.5) - 0.5 * log(0.5) &= 0.301 \\
H(Y) &= \sum_{y \in \mathcal{Y}} -p(y) log(p(y)) \\
&= -0.4 * log(0.4) - 0.45 * log(0.45) - 0.15 * log(0.15) &= 0.439
\end{aligned}$$

Recall:

$$\begin{aligned}
Pr[X, Y] &= Pr[X] * Pr[Y|X] \\
Pr[Y|X] &= \frac{Pr[X, Y]}{Pr[Y]}
\end{aligned}$$

Next, we will compute the conditional entropies:

$$\begin{aligned}
H(X|Y=a) &= \sum_{x \in \mathcal{X}} -p(x|y=a)log(p(x|y=a)) \\
&= -\frac{0.15}{0.5}log(\frac{0.15}{0.5}) - \frac{0.25}{0.5}log(\frac{0.25}{0.5}) \\
&= 0.307
\end{aligned}$$

$$\begin{aligned}
H(X|Y=b) &= \sum_{x \in \mathcal{X}} -p(x|y=b)log(p(x|y=b)) \\
&= -\frac{0.3}{0.5}log(\frac{0.3}{0.5}) - \frac{0.15}{0.5}log(\frac{0.15}{0.5}) \\
&= 0.290
\end{aligned}$$

$$\begin{aligned}
H(X|Y=c) &= \sum_{x \in \mathcal{X}} -p(x|y=c)log(p(x|y=c)) \\
&= -\frac{0.05}{0.5}log(\frac{0.05}{0.5}) - \frac{0.1}{0.5}log(\frac{0.1}{0.5}) \\
&= 0.240
\end{aligned}$$

$$\begin{aligned}
H(X|Y) &= \sum_{y \in \mathcal{Y}} p(y)H(X|Y=y) \\
&= 0.4 * 0.307 + 0.45 * 0.29 + 0.15 * 0.240 \\
&= .289
\end{aligned}$$

or, more concisely (ignoring rounding errors):

$$\begin{aligned}
H(X|Y) &= \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} p(x,y)log(\frac{p(y)}{p(x,y)}) \\
&= 0.15log(\frac{0.4}{0.15}) + 0.3log(\frac{0.45}{0.3}) + 0.05log(\frac{0.15}{0.05}) \\
&\quad + 0.25log(\frac{0.4}{0.25}) + 0.15log(\frac{0.45}{0.15}) + 0.1log(\frac{0.15}{0.1}) \\
&= 0.281 \\
H(Y|X) &= \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} p(x,y)log(\frac{p(x)}{p(x,y)}) \\
&= 0.15log(\frac{0.5}{0.15}) + 0.3log(\frac{0.5}{0.3}) + 0.05log(\frac{0.5}{0.05}) \\
&\quad + 0.25log(\frac{0.5}{0.25}) + 0.15log(\frac{0.5}{0.15}) + 0.1log(\frac{0.5}{0.1}) \\
&= 0.419
\end{aligned}$$

Thus, we obtain:

$$\begin{aligned}
H(X,Y) &= H(X) + H(Y|X) \\
&= 0.301 + 0.419 \\
&= 0.720 \\
H(X,Y) &= H(Y) + H(X|Y) \\
&= 0.439 + 0.281 \\
&= 0.720
\end{aligned}$$

Finally, we compute:

$$\begin{aligned}
I(X;Y) &= H(X) - H(X|Y) \\
&= 0.301 - 0.281 \\
&= 0.02 \\
I(Y;X) &= H(Y) - H(Y|X) \\
&= 0.439 - 0.419 \\
&= 0.02
\end{aligned}$$

$\square$

3 Question 3

In this question we given the distribution $p_\lambda = (1-\lambda) * p_0 + \lambda * p_1$, asked to prove relationships between $H(p_\lambda), H(p_0), H(p_1), \lambda$, and D.

The immediately obvious ones are:

$$
\begin{aligned}
H(p_\lambda) &= H(p_0) && (\lambda = 0)\\
H(p_\lambda) &= H(p_1) && (\lambda = 1)\\
H(p_\lambda) &\leq H(p_0) + H(p_1)
\end{aligned}
$$

$$D(p||q) = D(p||q) = 0 \Rightarrow p = q \Rightarrow H(p_\lambda) = H(p_0) = H(p_1)$$

Intuitively, it seems like something of the form

$$H(p_\lambda) = (1-\lambda) * H(p_0) + \lambda * H(p_1)$$

Note that

$$
\begin{aligned}
p_0 &= \frac{p_\lambda - \lambda p_1}{1-\lambda}\\
p_1 &= \frac{p_\lambda - (1-\lambda)p_0}{\lambda}
\end{aligned}
$$

Thus

$$
\begin{aligned}
H(p_1) &= \sum_{x \in \mathcal{X}} \frac{p_\lambda - (1-\lambda)p_0}{\lambda} log(\frac{p_\lambda - (1-\lambda)p_0}{\lambda}) \\
&= \sum_{x \in \mathcal{X}} \frac{p_\lambda - (1-\lambda)p_0}{\lambda} log(p_\lambda - (1-\lambda)p_0) - \frac{p_\lambda - (1-\lambda)p_0}{\lambda} log(\lambda) \\
&= \sum_{x \in \mathcal{X}} \frac{p_\lambda}{\lambda} log(p_\lambda - (1-\lambda)p_0) - \frac{1}{\lambda} log(p_\lambda - (1-\lambda)p_0) + \frac{\lambda p_0}{\lambda} log(p_\lambda - (1-\lambda)p_0) - \frac{p_\lambda - (1-\lambda)p}{\lambda} \\
&= \sum_{x \in \mathcal{X}} \frac{p_\lambda}{\lambda} log(\lambda p_1) - \frac{1}{\lambda} log(\lambda p_1) + \frac{\lambda p_0}{\lambda} log(\lambda p_1) - \frac{p_\lambda - (1-\lambda)p_0}{\lambda} log(\lambda) \\
&= \sum_{x \in \mathcal{X}} \frac{(1-\lambda)p_0 + \lambda p_1}{\lambda} log(\lambda p_1) - \frac{1}{\lambda} log(\lambda p_1) + \frac{\lambda p_0}{\lambda} log(\lambda p_1) - \frac{p_\lambda - (1-\lambda)p_0}{\lambda} log(\lambda) \\
&= \sum_{x \in \mathcal{X}} \frac{(1-\lambda)p_0}{\lambda} log(\lambda p_1) + p_1 log(\lambda p_1) - \frac{1}{\lambda} log(\lambda p_1) + \frac{\lambda p_0}{\lambda} log(\lambda p_1) - \frac{p_\lambda - (1-\lambda)p_0}{\lambda} log(\lambda)
\end{aligned}
$$

$$
\begin{aligned}
H(p_0) &= \sum_{x \in \mathcal{X}} \frac{p_\lambda - \lambda p_1}{1-\lambda} log(\frac{p_\lambda - \lambda p_1}{1-\lambda}) \\
&= \sum_{x \in \mathcal{X}} \frac{p_\lambda - \lambda p_1}{1-\lambda} log(p_\lambda - \lambda p_1) - \frac{p_\lambda - \lambda p_1}{1-\lambda} log(1-\lambda) \\
&= \sum_{x \in \mathcal{X}} \frac{p_\lambda}{1-\lambda} log(p_\lambda - \lambda p_1) - \frac{\lambda p_1}{1-\lambda} log(p_\lambda - \lambda p_1) - \frac{p_\lambda - \lambda p_1}{1-\lambda} log(1-\lambda) \\
&= \sum_{x \in \mathcal{X}} \frac{p_\lambda}{1-\lambda} log((1-\lambda)p_0) - \frac{\lambda p_1}{1-\lambda} log((1-\lambda)p_0) - \frac{p_\lambda - \lambda p_0}{1-\lambda} log(1-\lambda) \\
&= \sum_{x \in \mathcal{X}} \frac{(1-\lambda)p_0 + \lambda p_1}{1-\lambda} log((1-\lambda)p_0) - \frac{\lambda p_1}{1-\lambda} log((1-\lambda)p_0) - \frac{p_\lambda - \lambda p_0}{1-\lambda} log(1-\lambda) \\
&= \sum_{x \in \mathcal{X}} p_0 log((1-\lambda)p_0) + \frac{\lambda p_1}{1-\lambda} log((1-\lambda)p_0) - \frac{\lambda p_1}{1-\lambda} log((1-\lambda)p_0) - \frac{p_\lambda - \lambda p_0}{1-\lambda} log(1-\lambda)
\end{aligned}
$$

$$
\begin{aligned}
H(p_\lambda) &= \sum_{x \in \mathcal{X}} ((1-\lambda)p_0 + \lambda p_1) log((1-\lambda)p_0 + \lambda p_1) \\
&= \sum_{x \in \mathcal{X}} (1-\lambda)p_0 log((1-\lambda)p_0 + \lambda p_1) + \lambda p_1 log((1-\lambda)p_0 + \lambda p_1) \\
&= \sum_{x \in \mathcal{X}} log((1-\lambda)p_0 + \lambda p_1) - \lambda p_0 log((1-\lambda)p_0 + \lambda p_1) + \lambda p_1 log((1-\lambda)p_0 + \lambda p_1) \\
&= \sum_{x \in \mathcal{X}} log(p_\lambda) - \lambda p_0 log(p_\lambda) + \lambda p_1 log(p_\lambda)
\end{aligned}
$$

$$
\begin{aligned}
H(p_\lambda) - H(p_0) &= \sum_{x \in \mathcal{X}} log(p_\lambda) - \lambda p_0 log(p_\lambda) + \lambda p_1 log(p_\lambda) - p_0 log(p_0) \\
&= \sum_{x \in \mathcal{X}} log(p_\lambda) - p_0(\lambda log(p_\lambda) + log(p_0)) + \lambda p_1 log(p_\lambda)
\end{aligned}
$$

Nevertheless, after countless wasted hours and scratch paper, I am unable to come up with anything more interesting than the obvious w.r.t. these relationships...

4 Question 4

For this question, we were asked to prove or disprove the following relationship:

$$D(p||q) + D(q||r) \geq D(p||r)$$

We can disprove this with a counter example. First, write out our definitions.

$$D(p||q) = \sum_{x \in \mathcal{X}} p(x) log(\frac{p(x)}{q(x)})$$
$$D(q||r) = \sum_{x \in \mathcal{X}} q(x) log(\frac{q(x)}{r(x)})$$
$$D(p||r) = \sum_{x \in \mathcal{X}} p(x) log(\frac{p(x)}{r(x)})$$

Thus, we would like to show the following:

$$\sum_{x \in \mathcal{X}} p(x) log(\frac{p(x)}{q(x)}) + \sum_{x \in \mathcal{X}} q(x) log(\frac{q(x)}{r(x)}) \geq \sum_{x \in \mathcal{X}} p(x) log(\frac{p(x)}{r(x)}) \tag{8}$$
$$\sum_{x \in \mathcal{X}} p(x) log(\frac{p(x)}{q(x)}) - \sum_{x \in \mathcal{X}} p(x) log(\frac{p(x)}{r(x)}) \geq \sum_{x \in \mathcal{X}} q(x) log(\frac{q(x)}{r(x)}) \tag{9}$$
$$\sum_{x \in \mathcal{X}} p(x)[log(\frac{p(x)}{q(x)}) - log(\frac{p(x)}{r(x)})] \geq \sum_{x \in \mathcal{X}} q(x) log(\frac{q(x)}{r(x)}) \tag{10}$$
$$\sum_{x \in \mathcal{X}} p(x) log(\frac{r(x)}{q(x)}) \geq \sum_{x \in \mathcal{X}} q(x) log(\frac{q(x)}{r(x)}) \tag{11}$$
$$\sum_{x \in \mathcal{X}} [p(x) log(\frac{r(x)}{q(x)}) - q(x) log(\frac{q(x)}{r(x)})] \geq 0 \tag{12}$$

If we assume that $p(x) = q(x)$, this reduces to:

$$\sum_{x \in \mathcal{X}} p(x)[log(\frac{r(x)}{p(x)} - log(\frac{p(x)}{r(x)})] \geq 0 \tag{13}$$
$$\sum_{x \in \mathcal{X}} p(x) log(\frac{r(x)^2}{p(x)^2}) \geq 0 \tag{14}$$
$$\sum_{x \in \mathcal{X}} p(x) log[(\frac{r(x)}{p(x)})^2] \geq 0 \tag{15}$$
$$\sum_{x \in \mathcal{X}} 2p(x) log(\frac{r(x)}{p(x)}) \geq 0 \tag{16}$$
$$\tag{17}$$

If we let $\mathcal{X} = \{0, 1\}$ s.t. $p(0) = p(1) = .5$, and $r(0) = .7, r(1) = .3)$, then:

$$\sum_{x \in \mathcal{X}} 2p(x)log(\frac{r(x)}{p(x)}) = 2 * 0.5log(\frac{.7}{.5}) + 2 * 0.5log(\frac{.3}{.5})$$
$$= -0.076 < 0$$

Which violates our assumption, thus this inequality does not hold in the general case. $\square$