ECE401/8001
VISUAL SIGNAL PROCESSING
AND COMMUNICATIONS
Winter 2005
Prof.: Dr. Zhihai (Henry) He

Homework 01
01/27/2005

Luis M Vicente
Student number: 945-995
E-mail: lmvy8c@mizzou.edu

Problem 1 (25 points): Lets (X,Y) have the following joint distribution p(x,y)

y\x	0	1
0	1/4	1/8
1	1/2	1/8

Find

a) $H(X)$, $H(Y)$

b) $H(X/Y)$, $H(Y/X)$

c) $H(X,Y)$

d) $H(Y)-H(Y/X)$

e) $I(X,Y)$

Problem 2 (10 points): Let $\{x_1, x_2 ... x_n\}$ be arbitrary positive numbers, and be $\{a_1, a_2, ..., a_n\}$ be positive numbers whose sum is the unit. Prove that:

$$x_1^{a_1} x_2^{a_2} .... x_n^{a_n} \leq \sum_{i=1}^{n} a_i x_i$$

With equality if and only if all x_i are equal.

Problem 3 (15 points): Coin flips. A fair coin is flipped until the first head occurs. Let X the number of flips required. Find the entropy H(X). The following expression may be useful.

$$\sum_{n=1}^{\infty} r^n = \frac{r}{1-r}, \sum_{n=1}^{\infty} nr^n = \frac{r}{(1-r)^2}$$

Solution Problem 1 (25 points): Lets (X,Y) have the following joint distribution p(x,y)

		x	
	y\x	0	1
y	0	1/4	1/8
	1	1/2	1/8

Find

a) $H(X), H(Y)$

b) $H(X/Y), H(Y/X)$

c) $H(X,Y)$

d) $H(Y)-H(Y/X)$

e) $I(X,Y)$

Solution

a) $H(X)$**,** $H(Y)$

Entropy is defined as

$$H(X)=\sum_{x\in X} p(x)\log_2(\frac{1}{p(x)})$$

We can find the marginal distribution of X as:

$$p(x)=p(x,y=0)+p(x,y=1)$$

$$p(x=0)=p(0,0)+p(0,1)=\frac{1}{4}+\frac{1}{2}=\frac{3}{4}=0.75$$

$$p(x=1)=p(1,0)+p(1,1)=\frac{1}{8}+\frac{1}{8}=\frac{1}{4}=0.25$$

p(x)	3/4	1/4

Also

$$p(y)=p(x=0,y)+p(x=1,y)$$

$$p(y=0)=p(0,0)+p(1,0)=\frac{1}{4}+\frac{1}{8}=\frac{3}{8}=0.3750$$

$$p(y=1)=p(0,1)+p(1,1)=\frac{1}{2}+\frac{1}{8}=\frac{5}{8}=0.6250$$

p(y)	3/8	5/8

Therefore the Entropies for this exercise are:

$$H(X)=\sum_{x=0}^{1} p(x)\log_2(\frac{1}{p(x)})=p(x=0)\log_2(\frac{1}{p(x=0)})+p(x=1)\log_2(\frac{1}{p(x=1)})=$$

$$=\frac{3}{4}\log_2(\frac{4}{3})+\frac{1}{4}\log_2(4)=0.81128$$

$$H(Y)=\sum_{y=0}^{1} p(y)\log_2(\frac{1}{p(y)})=\frac{3}{8}\log_2(\frac{8}{3})+\frac{5}{8}\log_2(\frac{8}{5})=0.95443$$

b) $H(X/Y)$**,** $H(Y/X)$

Conditional Entropy; for this exercise can be written as:

$$H(X/Y) = -\sum_{x=0}^{1}\sum_{y=0}^{1} p(x,y)\log_2(p(x/y)) = -\sum_{x=0}^{1}\sum_{y=0}^{1} p(x,y)\log_2(\frac{p(x,y)}{p(y)})$$

$$= -p(x=0,y=0)\log_2(\frac{p(x=0,y=0)}{p(y=0)}) - p(x=1,y=0)\log_2(\frac{p(x=1,y=0)}{p(y=0)})$$

$$- p(x=0,y=1)\log_2(\frac{p(x=0,y=1)}{p(y=1)}) - p(x=1,y=1)\log_2(\frac{p(x=1,y=1)}{p(y=1)})$$

$$= -\frac{1}{4}\log_2(\frac{1/4}{3/8}) - \frac{1}{8}\log_2(\frac{1/8}{3/8}) - \frac{1}{2}\log_2(\frac{1/2}{5/8}) - \frac{1}{8}\log_2(\frac{1/8}{5/8}) = 0.7956$$

Similarly

$$H(Y/Y) = -1\sum_{x=0}\sum_{y=0}^{1} p(x,y)\log_2(p(y/x)) = -\sum_{x=0}^{1}\sum_{y=0}^{1} p(x,y)\log_2(\frac{p(x,y)}{p(x)})$$

$$= -\sum_{x=0}^{1}\sum_{y=0}^{1} p(x,y)\log_2(\frac{p(x,y)}{p(y)})$$

$$= -p(x=0,y=0)\log_2(\frac{p(x=0,y=0)}{p(x=0)}) - p(x=1,y=0)\log_2(\frac{p(x=1,y=0)}{p(x=1)})$$

$$- p(x=0,y=1)\log_2(\frac{p(x=0,y=1)}{p(x=0)}) - p(x=1,y=1)\log_2(\frac{p(x=1,y=1)}{p(x=1)})$$

$$= -\frac{1}{4}\log_2(\frac{1/4}{3/4}) - \frac{1}{8}\log_2(\frac{1/8}{1/4}) - \frac{1}{2}\log_2(\frac{1/2}{3/4}) - \frac{1}{8}\log_2(\frac{1/8}{1/4}) = 0.9387$$

c) $H(X,Y)$

Joint Entropy is defined as:

$$H(X,Y) = \sum_{x\in X}\sum_{y\in Y} p(x,y)\log_2(\frac{1}{p(x,y)})$$

For this particular problem:

$$H(X,Y) = \sum_{x=0}^{1}\sum_{y=0}^{1} p(x,y)\log_2(\frac{1}{p(x,y)}) =$$

$$= -p(x=0,y=0)\log_2(p(x=0,y=0)) - p(x=1,y=0)\log_2(p(x=1,y=0))$$

$$- p(x=0,y=1)\log_2(p(x=0,y=1)) - p(x=1,y=1)\log_2(p(x=1,y=1))$$

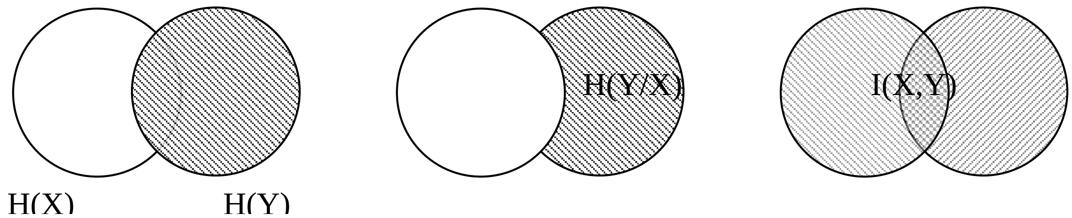
$$= -\frac{1}{4}\log_2(1/4) - \frac{1}{8}\log_2(1/8) - \frac{1}{2}\log_2(1/2) - \frac{1}{8}\log_2(1/8) = 1.75$$

d) $H(Y) - H(Y/X)$

For this part, let’s just plug in the values.

$H(Y) - H(Y/X) = 0.95443 - 0.93872 = 0.0157$

Let’s think the meaning of this equation. If we consider the Entropies as Venn diagrams we have that $H(Y) - H(Y/X)$ is just the part that intersects both diagrams. This is the mutual information.



e) $I(X,Y)$

Mutual Information is defined as

$$I(X,Y)=\sum_{x\in X}\sum_{y\in Y}p(x,y)\log_2\left(\frac{p(x,y)}{p(x)p(y)}\right)$$

For this particular problem:

$$I(X,Y)=\sum_{x=0}^{1}\sum_{y=0}^{1}p(x,y)\log_2\left(\frac{p(x,y)}{p(x)p(y)}\right)=$$

$$=p(x=0,y=0)\log_2\left(\frac{p(x=0,y=0)}{p(x=0)p(y=0)}\right)+p(x=1,y=0)\log_2\left(\frac{p(x=1,y=0)}{p(x=1)p(y=0)}\right)$$

$$+p(x=0,y=1)\log_2\left(\frac{p(x=0,y=1)}{p(x=0)p(y=1)}\right)+p(x=1,y=1)\log_2\left(\frac{p(x=1,y=1)}{p(x=1)p(y=1)}\right)$$

$$=\frac{1}{4}\log_2\left(\frac{1/4}{3/4*3/8}\right)+\frac{1}{8}\log_2\left(\frac{1/8}{1/4*3/8}\right)+\frac{1}{2}\log_2\left(\frac{1/2}{3/4*5/8}\right)+\frac{1}{8}\log_2\left(\frac{1/8}{1/4*5/8}\right)=0.0157$$

This result is in agreement with part d) because it is the same thing.

I have written a subroutine in Matlab that calculates all these results, regardless of the dimension and values of X and Y. Please find it attached at the APENDIX of this homework.

Solution Problem 2: Let $\{x_1, x_2 ... x_n\}$ be arbitrary positive numbers, and be $\{a_1, a_2, ..., a_n\}$ be positive numbers whose sum is the unit. Prove that:

$$x_1^{a_1} x_2^{a_2} .... x_n^{a_n} \leq \sum_{i=1}^{n} a_i x_i$$

with equality if and only if all x_i are equal.

Solution

If all x_i are equal we can arrange the equation to be

$$x_1^{a_1} x_1^{a_2} .... x_1^{a_n} \overset{?}{\leq} \sum_{i=1}^{n} a_i x_1$$

$$x_1^{a_1+a_2+..+a_n} \overset{?}{\leq} x_1 \sum_{i=1}^{n} a_i$$

$$x_1^{\sum_{i=1}^{n} a_i} \overset{?}{\leq} x_1 \sum_{i=1}^{n} a_i$$

Because $\sum_{i=1}^{n} a_i = 1$ we have

$$x_1^1 = x_1 * 1$$

If x_i are not equal

$$x_1^{a_1} x_2^{a_2} .... x_n^{a_n} \overset{?}{\leq} \sum_{i=1}^{n} a_i x_i$$ and we have that $\sum_{i=1}^{n} a_i = 1$

Jensen's inequality states that:

$$f\left(\sum_{i=1}^{N} p_i x_i\right) \leq \sum_{i=1}^{N} p_i f(x_i)$$ and $\sum_{i=1}^{N} p_i = 1$ being *f(x)* a convex function.

Let me rewrite the inequality as

$$f\left(\sum_{i=1}^{n} a_i x_i\right) \leq \sum_{i=1}^{n} a_i f(x_i)$$ and $\sum_{i=1}^{n} a_i = 1$

If we select a convex function like $f(x) = -\log_2(x)$ using Jensen's let me write:

$$-\log_2\left(\sum_{i=1}^{n} a_i x_i\right) \leq \sum_{i=1}^{n} a_i \left(-\log_2(x_i)\right)$$ and $\sum_{i=1}^{n} a_i = 1$

furthermore:

$$-\log_2\left(\sum_{i=1}^{n} a_i x_i\right) \leq \sum_{i=1}^{n} \left(-\log_2\left(x_i^{a_i}\right)\right)$$

$$-\log_2\left(\sum_{i=1}^{n} a_i x_i\right) \leq -\sum_{i=1}^{n} \log_2 x_i^{a_i}$$

Multiplying by -1 both sides we change the direction of the inequality as:

$$\log_2\left(\sum_{i=1}^{n} a_i x_i\right) \geq \sum_{i=1}^{n} \log_2 x_i^{a_i}$$

By the properties of logarithms we know that

$$\log_2 a + \log_2 b = \log_2(ab)$$

Therefore the right side can be expanded to the far right as:

$$\log_2\left(\sum_{i=1}^{n} a_i x_i\right) \geq \sum_{i=1}^{n} \log_2 x_i^{a_i} = \log_2(x_1^{a_1} x_2^{a_2} \ldots. x_n^{a_n})$$

Taking out the log from the left and the far right side we have the solution

$$\sum_{i=1}^{n} a_i x_i \geq x_1^{a_1} x_2^{a_2} \ldots. x_n^{a_n}$$

Solution Problem 3 (15 points): Coin flips. A fair coin is flipped until the first head occurs. Let X the number of flips required. Find the entropy H(X). The following expression may be useful.

$$\sum_{n=1}^{\infty} r^n = \frac{r}{1-r}, \sum_{n=1}^{\infty} nr^n = \frac{r}{(1-r)^2}$$

Solution

Let me write X as the RV to define the number of flips required. $x \in X\{x_1, x_2, x_3, ...\}$. In this case $p(x_1) = \frac{1}{2}$ is the probability to get a head at the first time.

$p(x_2) = \frac{1}{2} * \frac{1}{2} = \frac{1}{2^2}$ is the probability of getting tail at the first flip and getting head at the second flip.

$p(x_3) = \frac{1}{2} * \frac{1}{2} * \frac{1}{2} = \frac{1}{2^3}$ is the probability getting tails the first and the second flips, but get the head at the third flip. We can expand this reasoning until the *i* times

The Entropy is defined as:

$$H(X) = \sum_{x \in X} p(x) \log_2 (\frac{1}{p(x)})$$

In this case we have

$$H(X) = \sum_{i=1}^{\infty} p(x_i) \log_2 (\frac{1}{p(x_i)}) = \sum_{i=1}^{\infty} \frac{1}{2^i} \log_2 (2^i)$$

$$= \sum_{i=1}^{\infty} \frac{i}{2^i} \log_2 (2) = \sum_{i=1}^{\infty} \frac{i}{2^i} = \sum_{i=1}^{\infty} i \left(\frac{1}{2}\right)^i$$

I was given a useful expression $\sum_{n=1}^{\infty} nr^n = \frac{r}{(1-r)^2}$

Therefore:

$$H(X) = \sum_{i=1}^{\infty} i \left(\frac{1}{2}\right)^i = \frac{\frac{1}{2}}{\left(1-\frac{1}{2}\right)^2} = \frac{1/2}{1/4} = 2$$

APPENDIX

Problem 1 Appendix A

I have written a subroutine in Matlab that calculates all these results, regardless of the dimension and values of X and Y

This is the code:

```
%mprob2ass1
clear
clc
%This is the Probability matrix for X and for Y events.
pxy=[1/4, 1/8;
     1/2, 1/8];
%Size of alphabet
[N,N] = size(pxy);

%Marginal Probabilities. Be careful here, rows are y and columns are x
px(N,1)=0;
for ry=1:N %index of rows or y
    pi = pxy(ry,:)';
    px=px+pi;
end
%pxx = sum(pxy,1)

py(N,1)=0;
for cx=1:N %index of columns or x
    pi = pxy(:,cx);
    py=py+pi;
end
%pyy = sum(pxy,2)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Calculation of Entropy
HX = 0;
HY = 0;
for i=1:N %index for the marginal probabilities
    %to avoid dividing by zero
    if px(i)~=0
        HXi = px(i)*log2(1/px(i));
    else
        HXi = 0;
    end

    if py(i)~=0
        HYi = py(i)*log2(1/py(i));
    else
        HYi = 0;
    end

    HX = HX + HXi;
    HY = HY + HYi;
end
disp('Entropy from definition')
disp(['H(X) =  ',num2str(HX)]);
disp(['H(Y) =  ',num2str(HY)]);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Calculation of Joint Entropy
HXY = 0;
for ry=1:N %index for row/yvalues
    for cx=1:N %index for colum/xvalues
        %to avoid dividing by zero
        if pxy(ry,cx)~=0
            HXYi = pxy(ry,cx)*log2(pxy(ry,cx));
        else
            HXYi = 0;
        end
```

```
        HXY = HXY - HXYi;
    end
end
disp(['H(X,Y) =  ',num2str(HXY)]);


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Calculation of Conditional Entropy
%To do this, we have to find the Conditional probabilities of all elements.
%Lets calculate the matrix of conditional probabilities
for ry=1:N
    for cx=1:N
        pxcy(ry,cx) = pxy(ry,cx)/py(ry);
        pycx(ry,cx) = pxy(ry,cx)/px(cx); %I had to be very careful not
        %confuse rows with values of x (rows are values of y)
    end
end

%Now we calculate the Conditional Entropy H(X/Y) by its definition.
%We calculate it from the two formulas, and then verify that they
%agree with each other.
HXcY1 = 0;
HXcY2 = 0;
for ry=1:N
    for cx=1:N
        %to avoid dividing by zero
        if pxcy(ry,cx)~=0
            HXcYi1 = pxy(ry,cx)*log2(pxcy(ry,cx));
            HXcYi2 = pxy(ry,cx)*log2(pxy(ry,cx)/py(ry));
        else
            HXcYi1 = 0;
            HXcYi2 = 0
        end
        HXcY1 = HXcY1 - HXcYi1;
        HXcY2 = HXcY2 - HXcYi2;
    end
end

if HXcY1 ~= HXcY2
    disp('Error HXcY1 ~= HXcY2');
    disp(['H1(X/Y) =  ',num2str(HXcY1)]);
    disp(['H2(X/Y) =  ',num2str(HXcY2)]);
    return
else
    HXcY = HXcY2;
end
disp(['H(X/Y) =  ',num2str(HXcY)]);

%Now we will calculate the Conditional Entropy from the formula
disp(['H(X/Y) =  H(X,Y) - H(Y) = ',num2str(HXY-HY)]);

%Now we calculate the Conditional Entropy H(Y/X) by its definition.
%We calculate it from the two formulas, and then verify that they
%agree with each other.
HYcX1 = 0;
HYcX2 = 0;
for ry=1:N
    for cx=1:N
        %to avoid dividing by zero
        if pycx(ry,cx)~=0
            HYcXi1 = pxy(ry,cx)*log2(pycx(ry,cx));
            HYcXi2 = pxy(ry,cx)*log2(pxy(ry,cx)/px(cx));
        else
            HYcXi1 = 0;
            HYcXi2 = 0;
        end
        HYcX1 = HYcX1 - HYcXi1;
        HYcX2 = HYcX2 - HYcXi2;
    end
end
```

```
if HYcX1 ~= HYcX2
    disp('Error HYcX1 ~= HYcX2');
disp(['H1(Y/X) =  ',num2str(HYcX1)]);
disp(['H2(Y/X) =  ',num2str(HYcX2)]);
else
    HYcX = HYcX2;
end
disp(['H(Y/X) =  ',num2str(HYcX)]);
%Now we will calculate the conditional entropy from the formula
disp(['H(Y/X) =  H(X,Y) - H(X) = ',num2str(HXY-HX)]);

disp(' ');

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Calculation of Information
IXY = 0;

%From its definition
for ry=1:N
    for cx=1:N
        %to avoid dividing by zero
        if pxy(ry,cx)~=0 && px(cx)*py(ry)~=0
            IXYi = pxy(ry,cx)*log2(pxy(ry,cx)/(px(cx)*py(ry)));%I had to
            %be very careful not
            %confuse rows with values of x (rows are values of y)
        else
            IXYi = 0;
        end
        IXY = IXY + IXYi;
    end
end
disp(['I(X,Y) =  ',num2str(IXY)]);

%Now we will calculate the Information from the formula
disp(['I(X,Y) =  H(X) - H(X/Y) = ',num2str(HX-HXcY)]);
disp(['I(X,Y) =  H(Y) - H(Y/X) = ',num2str(HY-HYcX)]);

disp(['H(X) =  I(X,Y) + H(X/Y)= ',num2str(IXY+HXcY)]);
disp(['H(Y) =  I(X,Y) + H(Y/X)= ',num2str(IXY+HYcX)]);

disp(['H(X,Y) =  H(X) + H(Y) - I(X,Y) = ',num2str(HX+HY-IXY)]);
```

The screenshot output is

```
Entropy from definition
H(X) =  0.81128
H(Y) =  0.95443
H(X,Y) =  1.75
H(X/Y) =  0.79557
H(X/Y) =  H(X,Y) - H(Y) = 0.79557
H(Y/X) =  0.93872
H(Y/X) =  H(X,Y) - H(X) = 0.93872

I(X,Y) =  0.015712
I(X,Y) =  H(X) - H(X/Y) = 0.015712
I(X,Y) =  H(Y) - H(Y/X) = 0.015712
H(X) =  I(X,Y) + H(X/Y)= 0.81128
H(Y) =  I(X,Y) + H(Y/X)= 0.95443
H(X,Y) =  H(X) + H(Y) - I(X,Y) = 1.75
```