Sample paper 2008: two hour exam, do three questions. Solutions to q1 and 2.

1. (a) (9 marks) For discrete random variables X and Y define
   - The entropy $H(X)$.
   - The mutual information $I(X;Y)$.
   - The conditional entropy $H(X|Y)$.

   (b) (11 marks) Given the conditional distribution

	a	b	c
1	1/3	1/12	1/12
2	1/12	0	1/24
3	1/24	1/3	0

for $X \in \mathcal{X} = \{1,2,3\}$ and $Y \in \mathcal{Y} = \{a,b,c\}$, find $H(X)$, $H(Y)$, $H(X|Y)$, $H(Y|X)$ and $I(X;Y)$.

Solution: For a discrete random variable X with outcomes $\mathcal{X}$ and probability distribution $p_X(x)$ giving the probability of outcome $x \in \mathcal{X}$ the *entropy* is

$$H(X) = -\sum_{x\in\mathcal{X}} p_X(x) \log p_X(x) \tag{1}$$

If Y is another random variable, using the obvious notation for Y and joint distribution $p_{X,Y}(x,y)$, the *mutual information* between X and Y is

$$I(X;Y) = \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} p_{X,Y}(x,y) \log \frac{p_{X,Y}(x,y)}{p_X(x)p_Y(y)} \tag{2}$$

Finally, if $Y = y$ then the entropy of X is

$$H(X|Y=y) = -\sum_{x\in\mathcal{X}} p_{X|Y}(x|y) \log p_{X|Y}(x|y) \tag{3}$$

and the *conditional entropy* is the average of this

$$H(X|Y) = E_Y H(X|Y=y) = \sum_{y\in\mathcal{Y}} p(y) H(X|Y=y) \tag{4}$$

Now to work out $H(X)$ we need the marginal distribution for X

1	1/2
2	1/8
3	3/8

(5)

so

$$H(X) = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{8}\log\frac{1}{8} - \frac{3}{8}\log\frac{3}{8} \approx 1.4 \tag{6}$$

and to work out $H(Y)$ we need the marginal distribution for Y

$$\begin{array}{c|ccc} & \text{a} & \text{b} & \text{c} \\ \hline 1 & 11/24 & 5/12 & 3/24 \end{array} \tag{7}$$

so

$$H(Y) = -\frac{11}{24}\log\frac{11}{24} - \frac{5}{12}\log\frac{5}{12} - \frac{3}{24}\log\frac{3}{24} = 2 - \frac{3}{8}\log 3 \approx 1.41 \tag{8}$$

Now it begins to get annoying, working out the conditional probabilities,

$$\begin{aligned} H(X|Y=a) &= -\frac{8}{11}\log\frac{8}{11} - \frac{2}{11}\log\frac{2}{11} - \frac{1}{11}\log\frac{1}{11} \approx 1.09 \\ H(X|Y=b) &= -\frac{1}{5}\log\frac{1}{5} - \frac{4}{5}\log\frac{4}{5} \approx 0.72 \\ H(X|Y=c) &= -\frac{2}{3}\log\frac{2}{3} - \frac{1}{3}\log\frac{1}{3} \approx 0.92 \end{aligned} \tag{9}$$

and hence

$$H(X|Y) = \frac{11}{24}H(X|a) + \frac{5}{12}H(X|b) + \frac{3}{24}H(X|c) \approx 0.91 \tag{10}$$

The other way

$$\begin{aligned} H(Y|X=1) &= -\frac{2}{3}\log\frac{2}{3} - \frac{1}{3}\log\frac{1}{6} \approx 1.25 \\ H(Y|X=2) &= -\frac{2}{3}\log\frac{2}{3} - \frac{1}{3}\log\frac{1}{3} \approx 0.92 \\ H(Y|X=3) &= -\frac{1}{9}\log\frac{1}{9} - \frac{8}{9}\log\frac{8}{9} \approx 0.5 \end{aligned} \tag{11}$$

and hence

$$H(Y|X) = \frac{1}{2}H(Y|1) + \frac{1}{8}H(Y|2) + \frac{3}{8}H(Y|3) \approx 0.99 \tag{12}$$

Finally, we know

$$I(X;Y) = H(Y) - H(Y|X) = 1.41 - 0.99 = 0.42 \tag{13}$$

or,

$$I(X;Y) = H(X) - H(X|Y) = 1.40 - 0.91 = 0.49 \tag{14}$$

which reflects the rounding errors in the third digit.

2. For discrete random variables X, Y and Z

(a) (5 marks) prove
$$H(X,Y) = H(X) + H(Y|X)$$

(b) (5 marks) prove
$$I(X;Y) = H(Y) - H(Y|X)$$

(c) (5 marks) prove
$$H(X,Y|Z) \geq H(X|Z)$$

(d) (5 marks) prove
$$I(X;Z|Y) = I(Z;Y|X) - I(Z;Y) + I(X;Z)$$

Solution: The first two are from the book being Th 2.2.1 and part of Th 2.4.1; they basically follow from the definition and messing around with the logs and probabilities. The third one is just part one applied to relative entropy

$$\begin{aligned} H(X,Y|Z) &= E_Z H(X,Y|Z=z) \\ &= E_Z[H(X|Z=z) + H(Y|X,Z=z)] = H(X|Z) + H(Y|X,Z) \end{aligned} \tag{15}$$

and then
$$H(X,Y|Z) = H(X|Z) + H(Y|X,Z) \geq H(X|Z) \tag{16}$$

because $H(Y|X,Z) \geq 0$. Finally, part four is on a problem sheet, it is problem sheet 5, question 2d: by the chain rule for mutual information in different orders

$$\begin{aligned} I(X,Y;Z) &= I(X;Z|Y) + I(Z;Y) \\ I(X,Y;Z) &= I(Y;Z|X) + I(Z;X) \end{aligned} \tag{17}$$

so $I(X;Z|Y) + I(Z;Y) = I(Y;Z|X) + I(Z;X)$ and moving the $I(Z;Y)$ over the equals we get the equality.