18. For a subset α of $\mathcal{N}_n = \{1, 2, \cdots, n\}$, denote $(X_i, i \in \alpha)$ by X_α. For $1 \le k \le n$, let

$$H_k = \frac{1}{\binom{n}{k}} \sum_{\alpha : |\alpha| = k} \frac{H(X_\alpha)}{k}.$$

Here H_k is interpreted as the average entropy per random variable when k random variables are taken from $X_1, X_2, \cdots, X_n$ at a time. Prove that

$$H_1 \ge H_2 \ge \cdots \ge H_n.$$

This sequence of inequalities, due to Han [147], is a generalization of the independence bound for entropy (Theorem 2.39). See Problem 6 in Chapter 21 for an application of these inequalities.

Proof

First it can be proved that

$$H(X_{i1}X_{i2}\cdots X_{ik}) \le \frac{1}{k-1} \sum_{(j_1, j_2, \cdots j_{k-1}) \subset (i_1, i_2, \cdots i_k)} H(X_{j1}X_{j2}\cdots X_{jk-1}) \quad (1)$$

which is a generalization of Problem 17 in last homework and can be proved using similar method. Using (1) we have

$$\begin{aligned} H_k &= \frac{1}{\binom{n}{k}} \sum_{\alpha:|\alpha|=k} \frac{H(X_\alpha)}{k} \\ &\le \frac{1}{\binom{n}{k}} \frac{1}{k} \sum_{\alpha:|\alpha|=k} \frac{1}{k-1} \sum_{\beta\subset\alpha, \beta:|\beta|=k-1} H(X_\beta) \quad (2) \\ &= \frac{k!(n-k)!}{n!} \frac{1}{k} \frac{1}{k-1} \sum_{\alpha:|\alpha|=k} \sum_{\beta\subset\alpha, \beta:|\beta|=k-1} H(X_\beta) \end{aligned}$$

It can be proved that

$$\sum_{\alpha:|\alpha|=k} \sum_{\beta\subset\alpha, \beta:|\beta|=k-1} H(X_\beta) = (n-(k-1)) \sum_{\alpha:|\alpha|=k-1} H(X_\alpha) \quad (3)$$

because each $H(X_\alpha)$ is counted $(n-(k-1))$ times on the left side. Thus from (3) and (2), we have

$$\begin{aligned} & H_k \\ & \le \frac{k!(n-k)!}{n!} \frac{1}{k} \frac{1}{k-1} (n-k+1) \sum_{\alpha:|\alpha|=k-1} H(X_\alpha) \\ & = \frac{(k-1)!(n-k+1)!}{n!} \sum_{\alpha:|\alpha|=k-1} \frac{1}{k-1} H(X_\alpha) \quad (2) \\ & = H_{k-1} \end{aligned}$$

Q.E.D.

20. Prove the divergence inequality by using the log-sum inequality.

$$D(p//q)=\sum_{x\in S_p} p(x)\log\frac{p(x)}{q(x)}$$

$$\geq\left(\sum_{x\in S_p} p(x)\right)\cdot\log\frac{\sum_{x\in S_p} p(x)}{\sum_{x\in S_p} q(x)}\text{(according to log-sum inequality)}$$

0

1. Show that

$$I(X;Y;Z)=E\log\frac{p(X,Y)p(Y,Z)p(X,Z)}{p(X)p(Y)p(Z)p(X,Y,Z)}$$

and obtain a general formula for $I(X_1;X_2,;\cdots;X_n)$.

$$I(X;Y;Z)$$
$$=I(X;Y)-I(X;Y\mid Z)$$
$$=\sum_{x,y} p(x,y)\log\frac{p(x,y)}{p(x)p(y)}-\sum_{x,y,z} p(x,y,z)\log\frac{p(x,y\mid z)}{p(x\mid z)p(y\mid z)}$$
$$=\sum_{x,y,z} p(x,y,z)\log\frac{\frac{p(x,y)}{p(x)p(y)}}{\frac{p(x,y\mid z)}{p(x\mid z)p(y\mid z)}}$$
$$=\sum_{x,y,z} p(x,y,z)\log\frac{\frac{p(x,y)}{p(x)p(y)}}{\frac{p(x,y,z)p(z)p(z)}{p(z)p(x,z)p(y,z)}}$$
$$=\sum_{x,y,z} p(x,y,z)\log\frac{p(x,y)p(y,z)p(x,z)}{p(x,y,z)p(x)p(y)p(z)}$$
$$=E\log\frac{p(X,Y)p(Y,Z)p(X,Z)}{p(X,Y,Z)p(X)p(Y)p(Z)}$$

2. Suppose $X \perp Y$ and $X \perp Z$. Does $X \perp (Y, Z)$ hold in general?

Yes. It can be seen from the information diagram that I(X;YZ)=0

3. Show that $I(X;Y;Z)$ vanishes if at least one of the following conditions hold:

   a) X, Y, and Z are mutually independent;
   b) $X \to Y \to Z$ forms a Markov chain and X and Z are independent.

$$
\begin{aligned}
&I(X;Y;Z) \\
&= I(X;Y) - I(X;Y|Z) \\
&= H(X) + H(Y) - H(XY) - \left[H(X|Z) + H(Y|Z) - H(XY|Z)\right]
\end{aligned}
$$

if a) holds, i.e., X,Y and Z are mutually independent, we have

$$
\begin{aligned}
H(X|Z) &= H(X) \\
H(Y|Z) &= H(Y) \\
H(XY|Z) &= H(XY)
\end{aligned}
$$

So $I(X;Y;Z) = 0$

Also

$$
\begin{aligned}
&I(X;Y;Z) \\
&= I(X;Z) - I(X;Z|Y)
\end{aligned}
$$

If b) holds, i.e., X->Y->Z forms a Markov chain and X and Z are independent, we have

$$
\begin{aligned}
I(X;Z) &= 0 \\
I(X;Z|Y) &= 0
\end{aligned}
$$

4. a) Verify that $I(X;Y;Z)$ vanishes for the distribution $p(x, y, z)$ given by

$$p(0,0,0) = 0.0625,\ p(0,0,1) = 0.0772,\ p(0,1,0) = 0.0625,$$
$$p(0,1,1) = 0.0625,\ p(1,0,0) = 0.0625,\ p(1,0,1) = 0.1103,$$
$$p(1,1,0) = 0.1875,\ p(1,1,1) = 0.375.$$

b) Verify that the distribution in part (a) does not satisfy the conditions in Problem 3.

a) (X,Y) has the following distribution:

p(0,0)	p(0,1)	p(1,0)	p(1,1)
0.1397	0.125	0.1728	0.5625

I(X;Y)=0.054 bit

p(X,Y|Z=0) has the following distribution:

p(0,0)	p(0,1)	p(1,0)	p(1,1)
0.1667	0.1667	0.1667	0.5

I(X;Y|Z=0) = 0.045 bit

p(X,Y|Z=1) has the following distribution:

p(0,0)	p(0,1)	p(1,0)	p(1,1)
0.1667	0.1667	0.1667	0.5

I(X;Y|Z=0) = 0.0592 bit

I(X;Y|Z) is the expectation: I(X;Y|Z)=0.375*0.045+0.625*0.0592=0.054 bit

I(X;Y;Z) = I(X;Y) - I(X;Y|Z)= 0

b) p(X)

p(0)	p(1)
0.2647	0.7353

p(Y)

p(0)	p(1)
0.3125	0.6875

p(X)*p(Y)

p(0,0)	p(0,1)	p(1,0)	p(1,1)
0.0827	0.1820	0.2298	0.5055

So p(X)*p(Y) does not equal to p(X,Y). Thus X and Y are not independent.

(X,Z) the following distribution:

p(0,0)	p(0,1)	p(1,0)	p(1,1)
0.125	0.1397	0.25	0.4853

p(X)

p(0)	p(1)
0.2647	0.7353

P(Z)

p(0)	p(1)
0.375	0.625

p(X)*p(Z)

p(0,0)	p(0,1)	p(1,0)	p(1,1)
0.0993	0.1654	0.2757	0.4596

So p(X)*p(Z) does not equal to p(X,Z). Thus X and Z are not independent.

Both conditions in problem 3 do not hold.