P2.3 $H(\underline{p}) = -\sum_{l=1}^{n} p_l \log p_l$

and $-p_l \log p_l \geq 0$ for $0 \leq p_l \leq 1$

with equality when $p_l = 0$ or 1

So $H(\underline{p}) \geq 0$, and we have equality when $p_l = 0$ or 1 for all l.

P2.4 (a) Chain rule of entropy

(b) $H(g(X)|X) = 0$ since $P[g(X) = x_o | X] = 0$ or 1 for all x_o, and x.

(c) Chain rule of entropy

(d) Since $H(X|g(X)) \geq 0$.

P2.6 a) Let $X = 0$ or 1 with prob $1/2$ for each outcome.
Let $X = Y = Z$.
Then $I(X;Y|Z) = H(X|Z) - H(X|YZ)$

$$= 0 - 0$$
$$= 0$$

$$I(X;Y) = H(X) - H(X|Y)$$
$$= 1 - 0$$
$$= 1.$$

P 2.6 b) Let $X = 0$ or 1 with prob $1/2$ for each outcome.
Let $Y = 0$ or 1 with prob $1/2$ for each outcome, independently of X, i.e., $P_{XY} = P_X P_Y$.

Then
$$
\begin{aligned}
I(X;Y) &= D(P_{XY} \| P_X P_Y) \\
&= D(P_X P_Y \| P_X P_Y) \\
&= 0
\end{aligned}
$$

Now let $Z = X \oplus Y$, $\oplus$ is mod 2 addition

Then
$$
\begin{aligned}
I(X;Y|Z) &= I(X;Y|X \oplus Y) \\
&= H(X | X \oplus Y) - H(X | X \oplus Y, Y) \\
&= H(X) - H(X | X \oplus Y, Y, X) \\
&= 1 - 0 \\
&= \underline{1}.
\end{aligned}
$$

210 a) $H(X) = -\sum_{l=1}^{m} \alpha p_l \log(\alpha p_l)$

$$- \sum_{l=m+1}^{n} (1-\alpha) p_l \log\left((1-\alpha) p_l\right)$$

$$= -\alpha \sum_{l=1}^{m} p_l \left[\log p_l + \log \alpha\right]$$

$$-(1-\alpha) \sum_{l=m+1}^{n} p_l \left[\log p_l + \log(1-\alpha)\right]$$

$$= -\alpha \sum_{l=1}^{m} p_l \log p_l - \alpha \sum_{l=1}^{m} p_l \log \alpha$$

$$-(1-\alpha) \sum_{l=m+1}^{n} p_l \log p_l - (1-\alpha) \sum_{l=m+1}^{n} p_l \log(1-\alpha)$$

$$= \alpha H(X_1) + (1-\alpha) H(X_2)$$

$$- \alpha \log \alpha - (1-\alpha) \log(1-\alpha)$$

P2.11

a) $$\rho = 1 - \frac{H(X_2|X_1)}{H(X_1)}$$

$$= \frac{H(X_1) - H(X_2|X_1)}{H(X_1)}$$

$$= \frac{H(X_2) - H(X_2|X_1)}{H(X_1)}$$

$$= \frac{I(X_2;X_1)}{H(X_1)}$$

b) $$\frac{H(X_2|X_1)}{H(X_1)} = \frac{H(X_2|X_1)}{H(X_2)}$$

and this is between 0 and 1 since

$0 \le H(X_2|X_1) \le H(X_2)$.

c) When $H(X_2|X_1) = H(X_2) = H(X_1)$,

i.e., when X_2 is independent of X_1.

d) When $H(X_2|X_1) = 0$, ie, when X_2 is a function of X_1.

P2.14 a) $H(Z|X) = H(X+Y|X)$

$$= \sum_{\ell} P_X(\ell)\, H(Y + x_\ell \mid X = x_\ell)$$

$$= \sum_{\ell} P_X(\ell)\, H(Y \mid X = x_\ell)$$

$$= H(Y|X)$$

If X & Y are independent, then

$$H(Y) = H(Y|X) \quad \text{[independence used here]}$$
$$= H(Y+X|X)$$
$$\le H(Y+X) \quad \text{[unconditioning increases entropy]}$$
$$= H(Z)$$

b) Let $X = 0$ or 1 equally likely.
Let $Y = -X$
Then $H(X) = H(Y) = 1$
But $Z = X+Y = 0 \Rightarrow H(Z) = 0$

c) $H(X) + H(Y) \ge H(X,Y)$ (i)

$$= H(X,Y) + H(Z|X,Y) \quad \text{[Since } Z = X+Y\text{]}$$
$$= H(X,Y,Z)$$
$$= H(Z) + H(X,Y|Z)$$

If we also have that $H(X)+H(Y) = H(Z)$, Then, we have $H(Z) \geq H(Z) + H(X, Y \mid Z)$

Since $H(X, Y \mid Z) \geq 0$, this can only be true if

1) $H(X, Y \mid Z) = 0$

and 2) $H(X) + H(Y) = H(X, Y)$ in Step (i) above.

Condition 1) means that from the sum Z, we can recover both X & Y. This is true if for all $x_0, x_1 \in \mathcal{X}$, $y_0, y_1 \in \mathcal{Y}$,

$$x_0 + y_0 \neq x_1 + y_1$$

Condition 2) is true if X & Y are independent.

P2.15) $I(X_1; X_2, \dots, X_n)$

$$= I(X_1; X_2)$$
$$+ I(X_1; X_3 | X_2)$$
$$+ \cdots$$
$$+ I(X_1; X_n | X_{n-1}, \dots, X_2)$$

Let's look at a term of the form (for $\ell \geq 3$)

$$I(X_1; X_\ell | X_{\ell-1}, \dots, X_2)$$

$$= H(X_\ell | X_{\ell-1}, \dots, X_2) - H(X_\ell | X_{\ell-1}, \dots, X_2, X_1)$$

$$\overset{(a)}{=} H(X_\ell | X_{\ell-1}) - H(X_\ell | X_{\ell-1})$$

$$= 0$$

where (a) follows from Markov chain property.

Then $I(X_1; X_2, \dots, X_n) = I(X_1; X_2)$.

P2.17) (a) Since $X_1, \dots, X_n$ are iid with $H(X_\ell) = H(p)$ for $\ell = 1, \dots, n$.

(b) Since $(Z_1, \dots, Z_K, K)$ is a function of $(X_1, \dots, X_n)$ (see P2.4)

(c) chain rule of entropy

(d) Since conditioned on K, $Z_1, \dots, Z_K$ are iid Bernoullis with $H(Z_\ell) = 1$, $\ell = 1, \dots, n$

i.e.,

$$H(Z_1, \dots, Z_K | K) = \sum_{k=0}^{\infty} H(Z_1, \dots, Z_k | K = k)\, p(K=k)$$

$$= \sum_{k=0}^{\infty} k\, p(K=k)$$

$$= E[K]$$

(e) Since $H(K) \geq 0$

P 2.21) The "Markov inequality way":

$$P[p(X) \le d] = P\left[\frac{1}{p(X)} \ge \frac{1}{d}\right]$$

$$= P\left[\log \frac{1}{p(X)} \ge \log \frac{1}{d}\right]$$

$$\le \frac{E\left[\log \frac{1}{p(X)}\right]}{\log \frac{1}{d}} \quad \text{(Markov inequality)}$$

$$\Rightarrow \quad P[p(X) \le d] \log \frac{1}{d} \le H(X)$$

Alternate approach:

Let the pmf be given by $p_1 \le p_2 \le \cdots \le p_n$.

Assume that $p_l \le d < p_{l+1}$, i.e., p_l is last probability $\le d$.

Then
$$H(X) = \sum_{m=1}^{n} p_m \log \frac{1}{p_m}$$

$$\ge \sum_{m=1}^{l} p_l \log \frac{1}{p_l}$$

$$\ge \sum_{m=1}^{l} p_l \log \frac{1}{d}$$

$$= \left(\sum_{m=1}^{l} p_l\right) \log \frac{1}{d} = P[p(X) \le d] \log \frac{1}{d}$$

P2.27) Brute force calculation

$$H(q) + (p_{m-1} + p_m) H\left(\frac{p_{m-1}}{p_{m-1}+p_m}, \frac{p_m}{p_{m-1}+p_m}\right)$$

$$= \sum_{l=1}^{m-2} p_l \log \frac{1}{p_l} + \cancel{(p_{m-1}+p_m)}\left[\frac{p_{m-1}}{\cancel{p_{m-1}+p_m}} \log \frac{p_{m-1}+p_m}{p_{m-1}} + \frac{p_m}{\cancel{p_{m-1}+p_m}} \log \frac{p_{m-1}+p_m}{p_m}\right]$$

$$+ (p_{m-1}+p_m) \log\left[\frac{1}{(p_{m-1}+p_m)}\right]$$

$$= \sum_{l=1}^{m-2} p_l \log \frac{1}{p_l} + p_{m-1} \log \frac{1}{p_{m-1}} + p_m \log \frac{1}{p_m}$$

$$\cancel{+ p_{m-1} \log (p_{m-1}+p_m) + p_m \log (p_{m-1}+p_m)}$$

$$\cancel{+ (p_m + p_{m-1}) \log \frac{1}{p_{m-1}+p_m}}$$

$$= \sum_{l=1}^{m-2} p_l \log \frac{1}{p_l} + p_{m-1} \log \frac{1}{p_{m-1}} + p_m \log \frac{1}{p_m}$$

$$= H(p)$$

P2.28) Let $p = (p_1, \ldots, p_i, \ldots, p_j, \ldots, p_m)$

Let $q = (p_1, \ldots, \frac{p_i + p_j}{2}, \ldots, \frac{p_i + p_j}{2}, \ldots, p_m)$

Then $H(q) - H(p)$

$$= -\frac{p_i + p_j}{2} \log\left(\frac{p_i + p_j}{2}\right) - \frac{p_i + p_j}{2} \log\left(\frac{p_i + p_j}{2}\right) + p_i \log p_i + p_j \log p_j$$

$$= -(p_i + p_j) \log\left(\frac{p_i + p_j}{2}\right) + p_i \log p_i + p_j \log p_j$$

$$\geq 0 \qquad \text{with equality if } p_i = p_j$$

This is because by log-sum inequality:

$$p_i \log \frac{p_i}{1} + p_j \log \frac{p_j}{1} \geq (p_i + p_j) \log\left(\frac{p_i + p_j}{1+1}\right)$$

with equality if $\frac{p_i}{1} = \frac{p_j}{1} = \text{constant}$.

P2.29)

(a) $H(X,Y|Z) = H(X|Z) + H(Y|X,Z)$

$\geq H(X|Z)$

equality iff $H(Y|X,Z) = 0$ which is equivalent to Y is ~~independent~~ ~~of~~ a function of X & Z.

(b) $I(X,Y;Z) = I(X;Z) + I(Y;Z|X)$

$\geq I(X;Z)$

equality iff $I(Y;Z|X) = 0$ which is equivalent to Y & Z are independent given X.

(c) $H(X,Y,Z) - H(X,Y) \leq H(X,Z) - H(X)$ (1)

$\Leftrightarrow$ $H(Z|XY) \leq H(Z|X)$

but this is always true, hence (1) is always true.

Equality iff $H(Z|XY) = H(Z|X)$,

i.e., if $P_{Z|XY} = P_{Z|X}$

or equivalently $Y \to X \to Z$

(d) $I(X;Z|Y) \geq I(Z;Y|X) - I(Z;Y) + I(X;Z)$

$\iff I(X;Z|Y) + I(Z;Y) \geq I(Z;Y|X) + I(X;Z)$

$\iff I(X,Y;Z) \geq I(X,Y;Z)$

This is always true, and equality always holds.

P 2.31) $H(X|Y) = H(X|Y, g(Y))$

$\leq H(X|g(Y))$

with equality if $H(X|Y, g(Y)) = H(X|g(Y))$,

i.e., iff $H(X) - H(X|Y, g(Y)) = H(X) - H(X|g(Y))$

iff $I(X; Y, g(Y)) = I(X; g(Y))$

iff $I(X; Y, g(Y)) - I(X; g(Y)) = 0$

iff $I(X; Y|g(Y)) = 0$

iff $Y \to g(Y) \to X$.

P 2.34) Say $X_0 - X_1 - X_2 - \cdots - X_n - X_{n+1}$

Then $H(X_0 \mid X_n) = H(X_0 \mid X_n, X_{n+1})$

$$\leq H(X_0 \mid X_{n+1})$$

P 2.36) Let $p = (\frac{1}{3} \ \frac{2}{3})$

$q = (\frac{2}{3} \ \frac{1}{3})$

Then $D(p \| q) = \frac{1}{3} \log \frac{1/3}{2/3} + \frac{2}{3} \log \frac{2/3}{1/3}$

$$= D(q \| p)$$

P 2.37) $E\left[\log \frac{p(X,Y,Z)}{p(X)p(Y)p(Z)}\right]$

$$= E\left[\log \frac{1}{p(X)} + \log \frac{1}{p(Y)} + \log \frac{1}{p(Z)} - \log \frac{1}{p(X,Y,Z)}\right]$$

$$= H(X) + H(Y) + H(Z) - H(X,Y,Z)$$

$$\geq 0$$

with equality iff X & Y & Z are independent.

P2.39

a) $H(X,Y,Z) = H(X) + H(Y|\cancel{X}) + H(Z|XY)$

Now, $H(X) = H(Y) = H(Z) = 1$

Also, since $I(X;Y) = 0$, $H(Y) = H(Y|X)$

$\Rightarrow H(Y|X) = 1$

There are no constraints on $H(Z|XY)$, so let's pick it to be as small as possible, i.e., $=0$.

$\Rightarrow \quad H(X,Y,Z) \geq 1+1 = 2$

b) Let X, Y be iid Bernoulli$(1/2)$.

Let $Z = X + Y \mod 2$.

P2.40)

a) $H(X) = 8 \times \frac{1}{8} \log_2 \frac{1}{8} = 3$

b) $H(Y) = \sum_{l=1}^{\infty} 2^{-l} \log_2\left(\frac{1}{2^{-l}}\right)$

$$= \sum_{l=1}^{\infty} l\, 2^{-l}$$

$$= \sum_{l=0}^{\infty} l \left(\frac{1}{2}\right)^{l}$$

$$= \frac{1/2}{(1-1/2)^2} = 2$$

c) Because $(X+Y, X-Y)$ is a function of (X,Y), and (X,Y) is a function of $(X+Y, X-Y)$,

$$H(X+Y, X-Y) = H(X,Y)$$

$$= H(X) + H(Y)$$

$$= 3 + 2$$

$$= 5$$

P2.41) a) $I(X;Q,A)$

$$= \underbrace{I(X;Q)}_{\substack{=0 \\ \text{since } X \,\&\, Q \\ \text{are independent}}} + I(X;A|Q)$$

$$= H(A|Q) + \underbrace{H(A|Q,X)}_{=0 \text{ since } A = A(Q,X)}$$

$$= H(A|Q)$$

We interpret this as saying that, given the distribution of the questions, the more uncertainty there is in the answers, the more informative the answers are.

b) $I(X;Q_1 A_1 Q_2 A_2)$

$$= I(X;Q_1 A_1) + I(X;Q_2 A_2|Q_1 A_1)$$

$$= H(A_1|Q_1) + I(X;Q_2|Q_1 A_1) \}{=0}$$
$$+ I(X;A_2|Q_2 Q_1 A_1)$$

$$= H(A_1|Q_1) + 0 + H(A_2|Q_2 Q_1 A_1) + \underbrace{H(A_2|Q_2 Q_1 A_1 X)}_{=0}$$

$$= H(A_1|Q_1) + H(A_2|Q_2 Q_1 A_1)$$

$$\leq H(A_1|Q_1) + H(A_2|Q_2)$$

$$= H(A_1|Q_1) + H(A_1|Q_1) = 2I(X;Q_1 A_1)$$

P 2.42)

(a) Since $\cancel{X} = (5x) \div 5$

and $5x = (X) \cdot 5$

$H(X) = H(5X)$

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

$= H(Y) - H(Y|X, g(X))$

$\geq H(Y) - H(Y| g(X))$

$= I(g(X);Y)$

(c) $H(X_0|X_{-1}) \geq H(X_0|X_1, X_{-1})$

(d) $H(X,Y) = H(X) + H(Y|X)$

$\leq H(X) + H(Y)$

$\Rightarrow \dfrac{H(X,Y)}{H(X)+H(Y)} \leq 1$