1. $I(X^n; L^N) = H(X^n) - H(X^n | L^N)$

$$= n - \sum_{i=1}^{n} H(X_i | X^{i-1}, L^N)$$

and

$H(X_1 | L^N) = H(X_1)$ since $L^N$ and $X_1$ are independent

$H(X_i | X^{i-1}, L^N) = 0$ for $i = 2, 3, \ldots, n$

since $X_i$ is a function of $X_1$ and for $i = 2, 3, \ldots, n$

---

2. a. Conditioning on $X = x \in \mathcal{X}$, $\tilde{M} = M - 1$ is a geometric variable with probability $p(x)$, i.e.,

$$p_{\tilde{M}|X}(m|x) = p(x)\left(1 - p(x)\right)^{m-1} \quad , m \geq 1$$

$$\Rightarrow E[\tilde{M} | X = x] = \frac{1}{p(x)}$$ (longer waiting time for lower probability)

$$\Rightarrow E[\tilde{M}] = \sum_{x \in \mathcal{X}} p(x) E[\tilde{M} | X = x] = |\mathcal{X}|$$

b. $$E[\log \tilde{M}] = \sum_{x \in \mathcal{X}} p(x) E[\log \tilde{M} | X = x]$$

$$\leq \sum_{x \in \mathcal{X}} p(x) \log E[\tilde{M} | X = x]$$

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

3. $M_n = X_1 X_2 \dots X_n$

a. $P_{M_n}(m) = \begin{cases} \frac{1}{2^n} & \text{for } m = 2^n \\ 1 - \frac{1}{2^n} & \text{for } m = 0 \end{cases}$

b. $E[M_n] = 2^n \frac{1}{2^n} + 0\left(1 - \frac{1}{2^n}\right) = 1$

$H(M_n) = -\frac{1}{2^n} \log_2 \frac{1}{2^n} - \left(1 - \frac{1}{2^n}\right) \log_2 \left(1 - \frac{1}{2^n}\right) \to 0$ as $n \to \infty$

c. Shannon code:

$l_0 = \left\lceil -\log_2\left(1 - \frac{1}{2^n}\right) \right\rceil = 1$ if $n = 4$

$l_{2^n} = \left\lceil -\log_2 \frac{1}{2^n} \right\rceil = n = 4$

$R = \frac{1}{16} \cdot 4 + \left(1 - \frac{1}{16}\right) \cdot 1 = 1.1875$

$H(X) = 0{,}3373$

$\Rightarrow R - H(X) = 0{,}8502 < 1$

4. We need to identify

$X_1 X_2 X_3$

using a binary code. Therefore the average length of the binary code (= number of binary questions) is bounded as

$$H(X_1X_2X_3) \leq E[\#\text{ questions}] < H(X_1X_2X_3) + 1$$

where

$$H(X_1X_2X_3) = H(p_1) + H(p_2) + H(p_3)$$

---

5. $$-\frac{1}{5}\log_2 p(0,0,1,0,0) = -\frac{1}{5}\log_2\left(0.1\,(1-0.1)^4\right) = 0.7860$$

$$H(X) = H(0.1) = 0.469$$

$$\Rightarrow x^5 \notin A_{0.1}^{(5)}$$

The smallest value of $\epsilon$ is

$$0.7860 - 0.469 = 0.3170$$

6

a. No. For instance, the average

$$E[S_n] = E[X_1 + \cdots + X_n] = n\,E[X]$$

$$= n\left(\tfrac{1}{2}\cdot 1 + \tfrac{1}{2}\cdot 2\right) = \tfrac{3}{2}n$$

depends on time $n$.

b. $H(S_1,\dots,S_n) = H(X_1,\dots,X_n) = n\,H(\tfrac{1}{2}) = n$

c. Even though it is not stationary, we can write $H(\mathcal{X}) = \lim_{n\to\infty} \frac{n}{n} = 1$

---

7. a. $\mu_0 = \frac{1}{2} = P_Y(0)$

$\mu_1 = \frac{1}{2} = P_Y(1)$

b. $$H(\mathcal{X}) = \lim_{n\to\infty} \frac{1}{n} H(X_1,\dots,X_n) =$$

$$= \lim_{n\to\infty} \frac{1}{n} H(X_1,\dots,X_n,Y_1,\dots,Y_n)$$

$Y^n$ is a function of $X^n$

$$= \lim_{n\to\infty} \frac{1}{n} H(Y^n) + \lim_{n\to\infty} \frac{1}{n} \underbrace{H(X^n|Y^n)}_{n\cdot H(X|Y) = n\cdot 2}$$

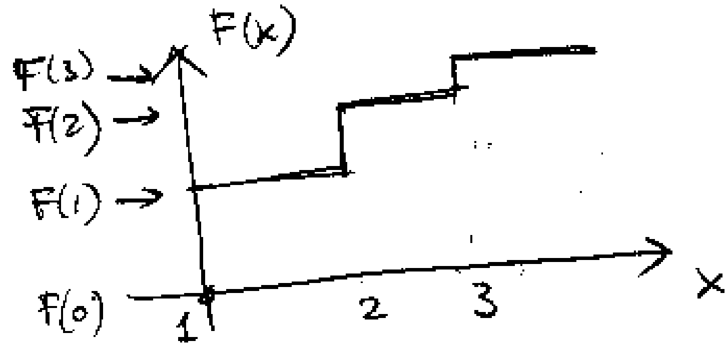
$$= H(\mathcal{Y}) + 2 = H(Y_2|Y_1) + 2$$

$$= \tfrac{1}{2}H(\tfrac{1}{3}) + \tfrac{1}{2}H(\tfrac{1}{3}) + 2$$

$$= H(\tfrac{1}{3}) + 2$$

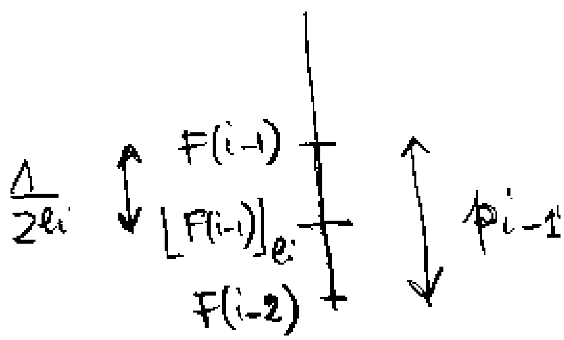
8.

a.




(m=3)

In order to show that the code is prefix-free, one must show that the intervals corresponding to the given code do not overlap




$$\frac{1}{2^{l_i}} = \frac{1}{2^{\lceil \log_2 p_i \rceil}} < \frac{1}{2^{-\log p_i}} = p_i \le p_{i-1} \quad \checkmark$$

b. Easy!

c. Easy!