

Information Theory HW1

許博翔

September 15, 2023

Problem 1. I'll prove (b) first, and then use (b) to prove (a) for convenience.

(b) Suppose that $s^n = (s_1, s_2, \dots, s_n) \in \mathcal{T}_\gamma^{(n)}(S)$.

By the definition of $\mathcal{T}_\gamma^{(n)}(S)$, $\forall a \in \mathbf{S}$, $\left| \frac{1}{n} \sum_{i=1}^n \mathbb{I}\{s_i = a\} - P_S(a) \right| \leq \gamma P_S(a)$.

$$\Rightarrow \forall a \in \mathbf{S}, \left| \frac{1}{n} \sum_{i=1}^n \log(P_S(a)) \mathbb{I}\{s_i = a\} - P_S(a) \log(P_S(a)) \right| \leq \gamma P_S(a) \log(P_S(a)).$$

$$\Rightarrow \sum_{a \in \mathbf{S}} \left| \frac{1}{n} \sum_{i=1}^n \log(P_S(a)) \mathbb{I}\{s_i = a\} - P_S(a) \log(P_S(a)) \right| \leq \sum_{a \in \mathbf{S}} \gamma P_S(a) \log(P_S(a)).$$

By triangular inequality,

$$\begin{aligned} & \left| \frac{1}{n} \sum_{i=1}^n \log(P_S(a)) + H(S) \right| \\ &= \left| \sum_{a \in \mathbf{S}} \frac{1}{n} \sum_{i=1}^n \log(P_S(a)) \mathbb{I}\{s_i = a\} - \sum_{a \in \mathbf{S}} P_S(a) \log(P_S(a)) \right| \\ &\leq \sum_{a \in \mathbf{S}} \left| \frac{1}{n} \sum_{i=1}^n \log(P_S(a)) \mathbb{I}\{s_i = a\} - P_S(a) \log(P_S(a)) \right| \\ &\leq \sum_{a \in \mathbf{S}} \gamma P_S(a) \log(P_S(a)) = -\gamma H(S). \end{aligned}$$

Taking $\delta = \xi(\gamma) := -\gamma H(S)$, and we get $\left| \frac{1}{n} \sum_{i=1}^n \log(P_S(a)) + H(S) \right| \leq \delta$, which

means $s^n \in \mathcal{A}_\delta^{(n)}(S)$.

$$\therefore \mathcal{T}_\gamma^{(n)}(S) \subseteq \mathcal{A}_\delta^{(n)}(S).$$

(a) Recall from (b), we take $\delta = \xi(\gamma) := -\gamma H(S)$.

The 4 properties in the proposition are:

(1) The original property is: $\forall s^n \in \mathcal{A}_\delta^{(n)}(S)$, $2^{-n(H(S)+\delta)} \leq \Pr\{S^n = s^n\} \leq 2^{-n(H(S)-\delta)}$.

\because from (b) we know that $\forall s^n \in \mathcal{T}_\gamma^{(n)}(S), s^n \in \mathcal{A}_\delta^{(n)}(S)$.

$$\therefore 2^{-n(H(S)+\delta)} \leq \Pr\{S^n = s^n\} \leq 2^{-n(H(S)-\delta)}.$$

(2) Let $A_n(a) := \{s^n \in \mathbf{S}^n : |\pi(a|s^n) - P_S(a)| > \gamma P_S(a)\}$.

Since $S \sim P_S$ is a DMS, the random variables $\{X_i\}_{i=1}^\infty$ where $X_i := \mathbb{I}\{S_i = a\}$ are i.i.d.

The average of X_i , denote as $\mu, = \Pr\{S_i = a\} = P_S(a)$.

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i = \frac{1}{n} \sum_{i=1}^n \mathbb{I}\{S_i = a\} = \pi(a|S^n).$$

Take $\epsilon > \gamma P_S(a)$.

By the weak law of large numbers, $\lim_{n \rightarrow \infty} \Pr\{S^n \in A_n(a)\} = \lim_{n \rightarrow \infty} \Pr\{|\pi(a|S^n) - P_S(a)| > \gamma P_S(a)\} = \lim_{n \rightarrow \infty} \Pr\{|\bar{X}_n - \mu| > \gamma P_S(a)\} \leq \lim_{n \rightarrow \infty} \Pr\{|\bar{X}_n - \mu| \geq \epsilon\} = 0$.

$$\therefore \mathcal{T}_\gamma^{(n)}(S) = \mathbf{S}^n \setminus \bigcup_{a \in \mathbf{S}} A_n(a).$$

$$\therefore \lim_{n \rightarrow \infty} \Pr\{S^n \in \mathcal{T}_\gamma^{(n)}(S)\} = 1 - \lim_{n \rightarrow \infty} \Pr\{S^n \in \bigcup_{a \in \mathbf{S}} A_n(a)\} \geq 1 - \lim_{n \rightarrow \infty} \sum_{a \in \mathbf{S}} \Pr\{S^n \in A_n(a)\} = 1.$$

$\therefore \forall \epsilon > 0$, by the definition of limits, $\Pr\{S^n \in \mathcal{T}_\gamma^{(n)}(S)\} \geq 1 - \epsilon$ for n large enough.

(3) $\because \mathcal{T}_\gamma^{(n)}(S) \subseteq \mathcal{A}_\delta^{(n)}(S)$.

$$\therefore |\mathcal{T}_\gamma^{(n)}(S)| \leq |\mathcal{A}_\delta^{(n)}(S)| \leq 2^{n(H(S)+\delta)}.$$

(4) By (2), $\forall \epsilon > 0$, for n large enough, there is $1 - \epsilon \leq \Pr\{S^n \in \mathcal{T}_\gamma^{(n)}(S)\} =$

$$\sum_{s^n \in \mathcal{T}_\gamma^{(n)}(S)} \Pr\{S^n = s^n\} \stackrel{(1)}{\leq} \sum_{s^n \in \mathcal{T}_\gamma^{(n)}(S)} 2^{-n(H(S)-\delta)} = |\mathcal{T}_\gamma^{(n)}(S)| 2^{-n(H(S)-\delta)}.$$

$$\therefore \forall \epsilon > 0, \text{ for } n \text{ large enough, there is } |\mathcal{T}_\gamma^{(n)}(S)| \geq (1 - \epsilon) 2^{n(H(S)-\delta)}.$$

Problem 2.

Problem 3.