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

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

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

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

By triangular inequality,

$$\left|\frac{1}{n}\sum_{i=1}^n \log(P_S(a)) + H(S)\right|$$

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

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

$$\le \sum_{a\in\mathbf{S}} \gamma P_S(a)\log(P_S(a)) = -\gamma H(S).$$

Taking $\delta = \xi(\gamma) := -\gamma H(S)$, and we get $\left|\frac{1}{n}\sum_{i=1}^n \log(P_S(a)) + H(S)\right| \le \delta$, 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: $\forall s^n \in \mathcal{A}_\delta^{(n)}(S)$, $2^{-n(H(S)+\delta)} \le \Pr\{S^n = s^n\} \le 2^{-n(H(S)-\delta)}$.

$\because$ from (b) we know that $\forall s^n \in \mathcal{T}_\gamma^{(n)}(S),\ s^n \in \mathcal{A}_\delta^{(n)}(S)$.
$\therefore 2^{-n(H(S)+\delta)} \le \Pr\{S^n = s^n\} \le 2^{-n(H(S)-\delta)}$.

(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 μ, $= \Pr\{S_i = a\} = P_S(a)$.
$\bar{X}_n = \dfrac{1}{n}\sum_{i=1}^{n} X_i = \dfrac{1}{n}\sum_{i=1}^{n} \mathbb{I}\{S_i = a\} = \pi(a|S^n)$.
Take $\epsilon > \gamma P_S(a)$.

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

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

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

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

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

(4) By (2), $\forall \epsilon > 0$, for n large enough, there is $1 - \epsilon \le \Pr\{S^n \in \mathcal{T}_\gamma^{(n)}(S)\} = \sum_{s^n \in \mathcal{T}_\gamma^{(n)}(S)} \Pr\{S^n = s^n\} \overset{(1)}{\le} \sum_{s^n \in \mathcal{T}_\gamma^{(n)}(S)} 2^{-n(H(S)-\delta)} = |\mathcal{T}_\gamma^{(n)}(S)| 2^{-n(H(S)-\delta)}$.
$\therefore \forall \epsilon > 0$, for n large enough, there is $|\mathcal{T}_\gamma^{(n)}(S)| \ge (1-\epsilon) 2^{n(H(S)-\delta)}$.

(c) Consider $\mathbf{S} = \{0, 1\},\ P_S(0) = P_S(1) = \dfrac{1}{2}, \gamma = 0.1$.

For the sequence $s^n = 0^n$, $|\pi(0|s^n) - P_S(0)| = \dfrac{1}{2} \not\le 0.05 = \gamma P_S(0)$.
$\Rightarrow 0^n \notin \mathcal{T}_\gamma^{(n)}(S)$.
However, $\forall \delta' > 0$, $\left| \dfrac{1}{n} \sum_{i=1}^{n} \log P_S(s_i) + H(S) \right| = \left| \dfrac{1}{n} \sum_{i=1}^{n} \log \dfrac{1}{2} - \log \dfrac{1}{2} \right| = 0 \le \delta'$.
$\Rightarrow 0^n \in \mathcal{A}_{\delta'}^{(n)}$.
$\therefore \mathcal{A}_{\delta'}^{(n)} \not\subseteq \mathcal{T}_\gamma^{(n)}(S)$.

Problem 2.

Problem 3.