

Information Theory HW1

許博翔

September 20, 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) $\therefore \mathcal{T}_\gamma^{(n)}(S) \subseteq \mathcal{A}_\delta^{(n)}(S)$.

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

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

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

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

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

For the sequence $s^n = 0^n$, $|\pi(0|s^n) - P_S(0)| = \frac{1}{2} \not\leq 0.05 = \gamma P_S(0)$.

$$\Rightarrow 0^n \notin \mathcal{T}_\gamma^{(n)}(S).$$

$$\text{However, } \forall \delta' > 0, \left| \frac{1}{n} \sum_{i=1}^n \log P_S(s_i) + H(S) \right| = \left| \frac{1}{n} \sum_{i=1}^n \log \frac{1}{2} - \log \frac{1}{2} \right| = 0 \leq \delta'.$$

$$\Rightarrow 0^n \in \mathcal{A}_{\delta'}^{(n)}.$$

$$\therefore \mathcal{A}_{\delta'}^{(n)} \not\subseteq \mathcal{T}_\gamma^{(n)}(S).$$

Problem 2.

(a) Define $X_i = \log \frac{1}{P_S(S_i)}$. Since S_i are i.i.d, X_i are also i.i.d.

Since $P_S(S_i) \leq 1$, we get that $\log \frac{1}{P_S(S_i)} \geq 0$.

$$\Rightarrow E[|X_i|] = E[X_i] = E[\log \frac{1}{P_S(S_i)}] = H(S) < \infty.$$

$$\prod_{i=1}^n P_S(S_i) \geq 2^{-n(H(S)+n^{-1/2}\delta_\zeta(S))}$$

$$\Leftrightarrow \prod_{i=1}^n \frac{1}{P_S(S_i)} \leq 2^{n(H(S)+n^{-1/2}\delta_\zeta(S))}$$

$$\Leftrightarrow \sum_{i=1}^n \log \frac{1}{P_S(S_i)} \leq n(H(S) + n^{-1/2}\delta_\zeta(S))$$

$$\Leftrightarrow \left(\frac{1}{n} \sum_{i=1}^n X_i \right) - H(S) \leq n^{-1/2}\delta_\zeta(S)$$

$$\Leftrightarrow \frac{\sqrt{n}(\bar{X}_n - E[X_i])}{\varsigma(S)} \leq \delta.$$

By central limit theorem, $\frac{\sqrt{n}(\bar{X}_n - E[X_i])}{\varsigma(S)} \xrightarrow{d} Z \sim N(0, 1)$ as $n \rightarrow \infty$.

$$\Rightarrow \Pr \left\{ \prod_{i=1}^n P_S(S_i) \geq 2^{-n(H(S)+n^{-1/2}\delta_\zeta(S))} \right\} = \Pr \left\{ \frac{\sqrt{n}(\bar{X}_n - E[X_i])}{\varsigma(S)} \leq \delta \right\}$$

$$\rightarrow \Pr\{Z \leq \delta\} = \Phi(\delta) \text{ as } n \rightarrow \infty.$$

(b) Since $\sum_{s^n \in \mathcal{B}_\delta^{(n)}(S)} P_{S^n}(s^n) \leq \sum_{s^n} P_{S^n}(s^n) = 1$,

and if $s^n \in B$, then $P_{S^n}(s^n) = \prod_{i=1}^n P_S(s_i) \geq 2^{-n(H(S)+n^{-1/2}\delta_\zeta(S))}$.

$$\therefore |\mathcal{B}_\delta^{(n)}(S)| 2^{-n(H(S)+n^{-1/2}\delta_\zeta(S))} = \sum_{s^n \in \mathcal{B}_\delta^{(n)}(S)} 2^{-n(H(S)+n^{-1/2}\delta_\zeta(S))} \leq \sum_{s^n \in \mathcal{B}_\delta^{(n)}(S)} P_{S^n}(s^n) \leq$$

1.

$$\Rightarrow |\mathcal{B}_\delta^{(n)}(S)| \leq 2^{n(H(S)+n^{-1/2}\delta_\zeta(S))}.$$

Let $Z \sim N(0, 1)$, by Berry-Esseen theorem, $|\Pr\{S^n \in \mathcal{B}_\delta^{(n)}(S)\} - \Pr\{Z \leq \delta\}| =$

$$\left| \Pr \left\{ \frac{\sqrt{n}(\bar{X}_n - E[X_i])}{\varsigma(S)} \leq \delta \right\} - \Pr\{Z \leq \delta\} \right| \leq cn^{-1/2} \text{ for some constant } c > 0.$$

$$\Rightarrow \Pr\{S^n \in \mathcal{B}_\delta^{(n)}(S)\} \geq \Pr\{Z \leq \delta\} - cn^{-1/2}.$$

Take $\delta = -\Phi^{-1}(\epsilon)$, we get that $\Phi(\delta) = 1 - \epsilon$.

$$\Rightarrow \Pr\{S^n \in \mathcal{B}_\delta^{(n)}(S)\} \rightarrow 1 - \epsilon \text{ as } n \rightarrow \infty.$$

Problem 3.

(a) Let $\delta \in (0, R - H(S))$, and $\mathcal{A}_\delta^{(n)}(S)$ be the δ -typical set defined in Definition 1.

By the third property of Proposition 1, we know that $|\mathcal{A}_\delta^{(n)}(S)| \leq 2^{n(H(S)+\delta)}$
 $H(S)+\delta < R \Rightarrow n(H(S)+\delta) < nR-1$ for n large enough
 $< 2^{\lfloor nR \rfloor}$ for n large enough.

$\Rightarrow \mathcal{A}_\delta^{(n)}(S)$ is an $(n, \lfloor nR \rfloor)$ code.

By the second property of Proposition 1, we know that $\forall \epsilon > 0, \exists N$ s.t. $\forall n \geq N, P_e^{(n)} = \Pr\{S^n \notin \mathcal{A}_\delta^{(n)}(S)\} \leq \epsilon$, which is the definition of $\lim_{n \rightarrow \infty} P_e^{(n)} = 0$.

\therefore such sequence exists, and it is $\mathcal{A}_\delta^{(n)}(S)$.