# Information-Theoretic Methods in Computer Science (CS1.502)

Spring 2023, IIIT Hyderabad
Assignment 2

Abhinav S Menon

## Problem 1

### Part (a)

We have $n$ i.i.d r.vs. $Z_1, \dots, Z_n$, each with mean $\mu$ and variance $\sigma^2$. We define

$$\overline{Z_n} = \frac{1}{n} \sum_{i=1}^{n} Z_i.$$

First, note that

$$\begin{aligned} \text{Var}(\overline{Z_n}) &= \text{Var}\left(\frac{1}{n}(Z_1 + \cdots + \overline{Z_n})\right) \\ &= \frac{1}{n^2}\text{Var}(Z_1 + \cdots + \overline{Z_n}) \\ &= \frac{n\sigma^2}{n^2} \\ &= \frac{\sigma^2}{n} \end{aligned}$$

and that

$$\mathbb{E}[\overline{Z_n}] = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[Z_i] = \mu.$$

Now, let us define a r.v. $X = (\overline{Z_n} - \mu)^2$. Then,

$$\begin{aligned}
\mathbb{E}[X] &= \int_0^\infty x P_X(x) dx \\
&= \int_0^{\varepsilon^2} x P_X(x) dx + \int_{\varepsilon^2}^\infty x P_X(x) dx \\
&\geq \varepsilon^2 \int_{\varepsilon^2}^\infty P_X(x) dx \\
&= \varepsilon^2 P_X(X \geq \varepsilon^2).
\end{aligned}$$

Substituting and rearranging, we see that

$$\begin{aligned}
P(|\overline{Z_n} - \mu| > \varepsilon) &= P((\overline{Z_n} - \mu)^2 > \varepsilon^2) \\
&\leq \frac{\sigma^2}{n\varepsilon^2},
\end{aligned}$$

which in the limit is zero. This allows to conclude that

$$\lim_{n\to\infty} P(|\overline{Z_n} - \mu| > \varepsilon) = 0,$$

QED.

## Part (b)

We are given a sequence $x^n = (x_1, \ldots, x_n)$ where $x_i \in \mathcal{X}$, and we define

$$\pi_{x^n}(x) = \frac{|i : x_i = x|}{n}.$$

Also, $X^n = (X_1, \ldots, X_n)$ is a sequence of i.i.d r.vs. distributed according to $P_X$ over $\mathcal{X}$.

Let us define, for any $x \in \mathcal{X}$, a set of $n$ random variables

$$Z_i = 1(x_i = x).$$

Then we have that $\overline{Z_n} = \pi_{x^n}(x)$, following the notation of Part (a) above.

Furthermore, the mean of each of the $Z_i$ (they are all the same because the variables are i.i.d) is $P_X(x)$.

Thus, applying the weak law of large numbers, we have

$$\lim_{n\to\infty} P(|\pi_{x^n}(x) > P_X(x)| > \varepsilon) = 0,$$

QED.

## Problem 2

We are given a tripartite graph with vertex set partitions $A, B, C$. We define $n_1$ as the number of edges between $A$ and $B$, $n_2$ as the number of edges between $B$ and $C$, and $n_3$ as the number of edges between $C$ and $A$. Furthermore, $n$ is the number of triangles in the graph.

Let $\mathcal{T}$ be the set of triangles in the graph. Further, let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be the sets of edges between $A$ and $B$, between $B$ and $C$, and between $C$ and $A$ respectively. Then we know that

$$\begin{aligned}
|\mathcal{T}| &\leq |\mathcal{X}||\mathcal{Y}||\mathcal{Z}| \\
|\mathcal{X}| &= n_1 \\
|\mathcal{Y}| &= n_2 \\
|\mathcal{Z}| &= n_3 \\
|\mathcal{T}| &= n
\end{aligned}$$

since a triangle is formed by one edge from each of the three sets, but not all such combinations form a triangle.

Now, let $T \sim \text{Unif}(\mathcal{T})$ be a triangle sampled from the graph. Similarly, let $X \sim \text{Unif}(\mathcal{X})$, $Y \sim \text{Unif}(\mathcal{Y})$, and $Z \sim \text{Unif}(\mathcal{Z})$ be edges sampled from each of the edge sets. Then, since

$$\begin{aligned}
H(X, Y, Z) &= H(X) + H(Y \mid X) + H(Z \mid X, Y) \\
H(X, Y) &= H(X) + H(Y \mid X) \\
H(Y, Z) &= H(Y) + H(Z \mid Y) \\
H(Z, X) &= H(X) + H(Z \mid X)
\end{aligned}$$

we can say that

$$\begin{aligned}
H(X, Y) + H(Y, Z) + H(Z, X) &= H(X) + H(X) \\
&\quad + H(Y \mid X) + H(Y) \\
&\quad + H(Z \mid Y) + H(Z \mid Y) \\
&\geq 2[H(X) + H(Y \mid X) + H(Z \mid X, Y)] \\
&= 2H(X, Y, Z).
\end{aligned}$$

This is Shearer's Lemma.

Now, we know that

$$H(X, Y, Z) = \log |\mathcal{X}||\mathcal{Y}||\mathcal{Z}|$$

and that

$$H(T) = \log |\mathcal{T}|,$$

since all the variables involved are independent and uniform. From the inequality we derived above, then, we can say that

$$H(T) \leq H(X, Y, Z).$$

This gives us

$$2H(T) \le H(X) + H(Y) + H(Z),$$

which gives us

$$2\log n \le n_1 + n_2 + n_3,$$

and so

$$n^2 \le n_1 n_2 n_3,$$

QED.

# Problem 3

We are given that $S$ is a r.v. drawn from $\mathcal{P}([1:n])$, and $P(i \in S) \ge \mu$ for all $i \in [1:n]$.
Let $S$ be drawn according to the distribution $P_S$. Then we can say that

$$P(i \in S) = \sum_{s \in \mathcal{P}([1:n])} P_S(s) \cdot 1(i \in S),$$

*i.e.*, the probability of $i$ being in $S$ is the sum of the probabilities of all sets including $i$.

Now,

$$\begin{aligned}
\mathbb{E}[H(X_S)] &= \sum_s P_S(s) H(X_S) \\
&\ge \sum_s P_S(s) \sum_{i \in s} H(X_i \mid X_{[1:i-1]}) \text{ [Shearer's Lemma]} \\
&= \sum_s P_S(s) \sum_{i=1}^{n} 1(i \in S) H(X_i \mid X_{[1:i-1]}) \\
&= \sum_{i=1}^{n} \left( \sum_S P_S(s) 1(i \in S) \right) H(X_i \mid X_{[1:i-1]}) \\
&\ge \sum_{i=1}^{n} \mu H(X_i \mid X_{[1:i-1]}) \\
&\ge \mu H(X_1, \ldots, X_n),
\end{aligned} .$$

QED.