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

Spring 2023, IIIT Hyderabad
Assignment 3

Abhinav S Menon

Problem 1

Part 1

We define, for an arbitrary permutation τ,

$$N_1(\sigma, \tau) = |\{1, 2, 3, 4\} - \{\sigma(\tau(1)), \sigma(\tau(2)), \dots \sigma(\tau(k-1))\}| = |[4] - P(k)|,$$

where $k = \tau^{-1}(1)$.
We need to show that for all τ such that $\tau(1) = j$, we have $N_1(\sigma, \tau) = 5 - j$.

We note first that if $\tau(1) = j$, then $j = \tau^{-1}(1) = k$.
Now, since σ and τ are both permutations, their composition $(\sigma \circ \tau)$ must be one also. Therefore $P(j)$, being a permutation of $[j-1]$, must have exactly $(j-1)$ elements.

Furthermore, since $P(j) \subseteq [4]$, we can say that

$$|[4] - P(j)| = 4 - |P(j)|.$$

Therefore $|N_1(\sigma, \tau)| = 4 - (j-1) = 5 - j$, QED.

Part 2

We have shown in Part 1 that

$$|N_1(\sigma, \tau)| = 5 - \tau^{-1}(1).$$

Therefore, all τ such that $|N_1(\sigma, \tau)| = j$ are such that $\tau^{-1}(1) = 5 - j$, so $1 = \tau(5 - j)$.

Therefore, the mapping of $5-j$ is fixed, and so it is left to map the set $[4]-\{5-j\}$ to $\{2, 3, 4\}$. The number of such permutations is $|S_3| = 3! = 6$, QED.

Question 2

Let $Y \sim \mathcal{N}(\mu, \sigma^2)$. Since entropy is independent of the mean, we can set $\mu = 0$ WLOG.
Let $X \sim P_X$ be an arbitrary r.v. with variance σ^2.

Now, note that

$$\begin{aligned} D(X \,||\, Y) &= \int_{\mathbb{R}} p_X(t) \log \frac{p_X(t)}{p_Y(t)} dt \\ &= \int_{\mathbb{R}} p_X(t) \log p_X(t) dt - \int_{\mathbb{R}} p_X(t) \log p_Y(t) dt \\ &= -h(X) - I. \end{aligned}$$

The value of I can be computed as follows:

$$\begin{aligned} I &= \int_{\mathbb{R}} p_X(t) \log p_Y(t) dt \\ &= \int_{\mathbb{R}} p_X(t) \log \left(\frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{t^2}{2\sigma^2}} \right) dt \\ &= \int_{\mathbb{R}} p_X(t) \log \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right) dt + \int_{\mathbb{R}} p_X(t) \log e^{-\frac{t^2}{2\sigma^2}} dt \\ &= -\frac{1}{2} \log(2\pi\sigma^2) \int_{\mathbb{R}} p_X(t) dt - \frac{\log e}{2\sigma^2} \int_{\mathbb{R}} t^2 p_X(t) dt \\ &= -\frac{1}{2} \log(2\pi\sigma^2) \cdot 1 - \frac{1}{2} \log e \\ &= -\frac{1}{2} \log(2\pi e \sigma^2). \end{aligned}$$

Therefore

$$D(X \,||\, Y) = -h(X) + \frac{1}{2} \log(2\pi e \sigma^2).$$

Since KL-divergence is always nonnegative, we can conclude that

$$h(X) \leq \frac{1}{2} \log(2\pi e \sigma^2),$$

QED.

Question 3

Consider the case where P_X and Q_X are binary distributions. Let $p = P_X(0)$ and $q = Q_X(0)$. Furthermore, let $p = \frac{1}{2}$ and $q = \frac{1}{2} + \varepsilon$.

Now, we have that

$$D_2(P_X \,||\, Q_X) \geq \frac{k}{\ln 2} \mathrm{TV}(P_X, Q_X)^2.$$

From this, we can say that

$$g_p(q) = D_2(P_X \mid\mid Q_X) - \frac{k}{\ln 2}\mathrm{TV}(P_X, Q_X)^2 \geq 0,$$

i.e., 0 is a minimum of $g_p(q)$, and so $\frac{d}{dq}g_p(q) = 0$.

Now,

$$\begin{aligned}
g_p'(q) = 0 &\implies \frac{p}{\ln 2}\log\frac{p}{q} + \frac{1-p}{\ln 2}\log\frac{1-p}{1-q} - \frac{k}{\ln 2}(p-q)^2 = 0 \\
&\implies -\frac{p}{q} + (1-p)\frac{1}{1-q} + 2k(p-q) = 0 \\
&\implies k = \frac{1}{2(p-q)}\left(\frac{p}{q} - (1-p)\frac{1}{1-q}\right) \\
&\implies k = \frac{1}{2(p-q)}\left(\frac{p(1-q) - q(1-p)}{q(1-q)}\right) \\
&\implies k = \frac{1}{2(p-q)}\left(\frac{p-q}{q(1-q)}\right) \\
&\implies k = \frac{1}{2q(1-q)}
\end{aligned}$$

WLOG, we can say that $q \leq \frac{1}{2}$. Therefore,

$$\begin{aligned}
q(1-q) &\leq \frac{1}{4} \\
\implies \frac{1}{q(1-q)} &\geq 4 \\
\implies k &\geq 2,
\end{aligned}$$

QED.

Question 4

We have two probability distributions P_X, Q_X and a function $f : \mathcal{X} \to [0, B]$. We need to show that

$$|\mathbb{E}_{P_X}[f(X)] - \mathbb{E}_{Q_X}[f(X)]| \leq B \cdot \mathrm{TV}(P_X, Q_X).$$

We can proceed as follows.

$$
\begin{aligned}
|\mathbb{E}_{P_X}[f(X)] - \mathbb{E}_{Q_X}[f(X)]| &= \left|\sum_{\mathcal{X}} P_X(x)f(x) - \sum_{\mathcal{X}} Q_X(x)f(x)\right| \\
&= \left|\sum_{\mathcal{X}} f(x)(P_X(x) - Q_X(x))\right| \\
&= \left|\sum_{\mathcal{X}}(P_X(x) - Q_X(x))\left(f(x) - \frac{B}{2}\right) + \sum_{\mathcal{X}}(P_X(x) - Q_X(x)) \cdot \frac{B}{2}\right| \\
&= \left|\sum_{\mathcal{X}}(P_X(x) - Q_X(x))\left(f(x) - \frac{B}{2}\right) + 0 \cdot \frac{B}{2}\right| \\
&= \left|\sum_{P_X(x) > Q_X(x)}(P_X(x) - Q_X(x))\left(f(x) - \frac{B}{2}\right)\right. \\
&\quad \left. - \sum_{P_X(x) \leq Q_X(x)}(Q_X(x) - P_X(x))\left(f(x) - \frac{B}{2}\right)\right| \\
&\leq \left|\sum_{P_X(x) > Q_X(x)}(P_X(x) - Q_X(x))\left(f(x) - \frac{B}{2}\right)\right| \\
&\quad + \left|\sum_{P_X(x) \leq Q_X(x)}(Q_X(x) - P_X(x))\left(f(x) - \frac{B}{2}\right)\right| \\
&= \left|f(x) - \frac{B}{2}\right| \left|\sum_{\mathcal{X}} P_X(x) - Q_X(x)\right| \\
&\leq \frac{B}{2}\left|\sum_{\mathcal{X}} P_X(x) - Q_X(x)\right| \\
&= B \cdot \mathrm{TV}(P_X, Q_X),
\end{aligned}
$$

QED.