CAS751: IT Metods in Trustworthy ML Homework 1 - **Due: 10/6/2024**

Information-Theoretic Measures

Instructor: Shahab Asoodeh

1. *(TV distance and Hockey-stick divergence. 30 points)* Let P and Q are two continuous distributions with pdf's p and q, respectively. Show that for $\gamma \geq 1$

   (a) $\mathsf{TV}(P, Q) = 1 - \int \min\{p(x), q(x)\}\mathrm{d}x$

   (b) $\mathsf{TV}(P, Q) = \int \max\{p(x), q(x)\}\mathrm{d}x$ -1

   (c) $\mathsf{E}_\gamma(P\|Q) = 1 - \int \min\{p(x), \gamma q(x)\}\mathrm{d}x$

   (d) $\mathsf{E}_\gamma(P\|Q) = \int \max\{p(x), \gamma q(x)\}\mathrm{d}x - \gamma$

   (e) $\mathsf{E}_\gamma(P\|Q) = \frac{1}{2}\int |p(x) - \gamma q(x)|\mathrm{d}x - \frac{1}{2}(\gamma - 1)$

2. *(Rényi divergences. 10 points)* Let $P = \prod_{i=1}^n P_i$ and $Q = \prod_{i=1}^n Q_i$, where $P_i = \mathcal{N}(\mu_i, \sigma^2)$ and $Q_i = \mathcal{N}(\nu_i, \sigma^2)$ for $i \in \{1, 2, \dots, n\}$. Compute $D_\alpha(P\|Q)$ for all $\alpha \in (0, 1) \cup (1, \infty)$.

3. *(Hockey-stick divergences. 20 points)*

   (a) Let $P_{Y|X}$ be defined as follows: $P_{Y|X=0} = \mathsf{Bernoulli}(1 - \delta)$ and $P_{Y|X=1} = \mathsf{Bernoulli}(\delta)$, where $\delta = \frac{1}{1+e^\varepsilon}$ for some $\varepsilon \geq 0$. Show that $\eta_\gamma(P_{Y|X}) = 0$ for $\gamma = e^\varepsilon$.

   (b) Let $\mathcal{A} = \{1, 2, \dots, k\}$ and $\gamma \geq 1$. Let $P_{Y|X}$ be defined as

   $$P_{Y|X}(y|x) = \begin{cases} \frac{\gamma}{\gamma+k-1}, & \text{if } x, y \in \mathcal{A}, x = y, \\ \frac{1}{\gamma+k-1}, & \text{if } x, y \in \mathcal{A}, x \neq y, \\ 0 & \text{otherwise.} \end{cases}$$

   Let P_X and Q_X be two arbitrary distributions on $\mathcal{A}$, and P_Y and Q_Y be the corresponding distributions induced by $P_{Y|X}$, respectively. Show that $P_Y(A) \leq \gamma Q_Y(A)$ for any $A \subseteq \mathcal{A}$.

4. *(Maximal coupling and TV. 20 points)* Let P and Q be two discrete distributions supported on a finite set $\mathcal{A}$. Show that

   $$\mathsf{TV}(P, Q) = \inf_{P_{XY}} \Pr(X \neq Y),$$

   where the infimum is taken over all joint distributions P_{XY} such that $\sum_y P_{XY}(x, y) = P(x)$ for any $x \in \mathcal{A}$ and $\sum_x P_{XY}(x, y) = Q(y)$ for any $y \in \mathcal{A}$. [Such P_{XY} are called *couplings* of P and Q.]

5. *(Strong DPI. 20 points)* Fix a channel $P_{Y|X}$. We defined in the class

   $$\eta_{\mathsf{TV}}(P_{Y|X}) \coloneqq \sup_{P,Q:P\neq Q} \frac{\mathsf{TV}(P_Y, Q_Y)}{\mathsf{TV}(P_X, Q_X)},$$

   where P_Y and Q_Y are the output distributions of $P_{Y|X}$ corresponding to input distributions P_X and Q_X, respectively. We then showed that

   $$\eta_{\mathsf{TV}}(P_{Y|X}) = \sup_{x_1, x_2 \in \mathcal{X}} \mathsf{TV}(P_{Y|X=x_1}, P_{Y|X=x_2}).$$

   In this problem, we wish to generalize this result to the hockey-stick divergence. Define for $\gamma \geq 1$

   $$\eta_\gamma(P_{Y|X}) \coloneqq \sup_{P,Q:P\neq Q} \frac{\mathsf{E}_\gamma(P_Y\|Q_Y)}{\mathsf{E}_\gamma(P_X\|Q_X)}.$$

   Show that

   $$\eta_\gamma(P_{Y|X}) = \sup_{x_1, x_2 \in \mathcal{X}} \mathsf{E}_\gamma(P_{Y|X=x_1}\|P_{Y|X=x_2}).$$