Problem Set 7

Issued: Wednesday 13th December, 2017 **Due:** Monday 18th December, 2017

7.1. Let $q(y) > 0$ $(y = 0, 1, \cdots)$ be a probability mass function for a random variable y and let $\mathcal{P}$ be the set of all PMFs defined over $\{0, \cdots, M-1\}$ for a known constant M:

$$\mathcal{P} \triangleq \{p(\cdot) | p(y) = 0 \text{ for all } y \geq M\}$$

We can represent each element p of $\mathcal{P}$ as a M-dimensional vector $[p_0, \cdots, p_{M-1}]^{\mathrm{T}}$ that lies on a $(M-1)$-dimensional simplex, i.e., $\sum_{m=0}^{M-1} p_m = 1$.

(a) Show that, for all $p \in \mathcal{P}$, $D(q \| p) = \infty$.

(b) Show that, for all $p \in \mathcal{P}$, $D(p \| q) < \infty$.

(c) Find the I-projection of q onto $\mathcal{P}$, $p^* = \arg\min_p D(p \| q)$, and the corresponding divergence $D(p^* \| q)$ in terms of $Q(y) \triangleq \mathbb{P}(\mathsf{y} \leq y)$, the CDF of the random variable y.

Let $\mathcal{P}_\epsilon$ be the space of all PMFs with weight of ϵ on values M and above:

$$\mathcal{P}_\epsilon \triangleq \left\{ p(\cdot) \middle| \sum_{y=M}^{\infty} p(y) = \epsilon \right\}$$

We can think of $\mathcal{P}_\epsilon$ as an extension of $\mathcal{P}$ to the distributions defined for all integers that only allows limited weight to be allocated to the values outside $\{0, \cdots, M-1\}$.

(d) Find the I-projection of q onto $\mathcal{P}_\epsilon$, $p_\epsilon^* = \arg\min_p D(p \| q)$, and the corresponding divergence $D(p_\epsilon^* \| q)$ in terms of $Q(y)$. Show that $\lim_{\epsilon \to 0^+} D(p_\epsilon^* \| q) = D(p^* \| q)$.

(e) Show that $\mathcal{P}_\epsilon$ can be represented as a linear family of PMFs, i.e.,

$$\mathcal{P}_\epsilon = \{p(\cdot) | \, \mathbb{E}_p[t(\mathsf{y})] = c\},$$

and invent the appropriate statistic $t(\cdot)$ and constant c.

(f) Show that p_ϵ^* belongs to the exponential family $\mathcal{E}(x, \lambda(x) = x, t(\cdot), \ln q(\cdot))$ and find the value of the parameter x that corresponds to p_ϵ^*.

7.2. Let x and y be discrete random variables with a joint distribution $p_{\mathsf{x},\mathsf{y}}(x, y)$. We wish to approximate this distribution with a separable distribution $q(x, y) = q_{\mathsf{x}}(x) q_{\mathsf{y}}(y)$.

(a) Find expressions for $q_{\mathsf{x}}(x)$ and $q_{\mathsf{y}}(y)$ that minimize $D(p_{\mathsf{x},\mathsf{y}} \,\|\, q)$.

(b) Say that x and y take on values in $\{1, 2, 3, 4\}$ and have the joint distribution

$$p_{\mathsf{x},\mathsf{y}}(x, y) = \begin{cases} \frac{1}{4} & x = y = 3 \text{ or } x = y = 4 \\ \frac{1}{8} & x \leq 2 \text{ and } y \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

Find the $q_{\mathsf{x}}(x)$ and $q_{\mathsf{y}}(y)$ that minimize $D(q \,\|\, p_{\mathsf{x},\mathsf{y}})$. Does your answer from part (a) give a value of $D(q \,\|\, p_{\mathsf{x},\mathsf{y}})$ close to the minimum?