Third homework set, Due April 17, 16:00

1. (1p.) Consider the following family of candidate distributions on $\mathcal{X} = \{1, \ldots, k\}$: the distributions of form $\mathbb{P}(x) = c \cdot \exp(t_1 x + t_2 x^2)$. Given a sample $\mathbf{x} = (x_1, \ldots, x_n)$, denote by $\mathbb{P}^*$ the maximum likelihood estimate provided that it exists. Assume that each symbol of $\mathcal{X}$ occurs in the sample $\mathbf{x}$. Show that the maximum likelihood estimate exists in this case. Specify linear set $\mathcal{L}$ of distributions on $\mathcal{X}$ such that $\mathbb{P}^*$ is equal to the I-projection of the uniform distribution onto $\mathcal{L}$.

2. (3p.) Let $\mathcal{E}$ be the family of binomial distributions with $n = 5$ and $p \in (0, 1)$, i.e.,

$$\mathcal{E} = \{\mathbb{P} : \mathbb{P}(a) = \binom{5}{a} p^a (1-p)^{5-a}, a \in \{0, 1, 2, 3, 4, 5\}, \text{ for some } p \in (0, 1).\} \tag{1}$$

   (a) Show that $\mathcal{E}$ is an exponential family!

   (b) We observe 200 independent drawing from an unknown distribution on $A = \{0, 1, 2, 3, 4, 5\}$. The type of the observed sample $\hat{\mathbb{P}}_{200} = (\hat{\mathbb{P}}_{200}(0), \hat{\mathbb{P}}_{200}(1), \hat{\mathbb{P}}_{200}(2), \hat{\mathbb{P}}_{200}(3), \hat{\mathbb{P}}_{200}(4), \hat{\mathbb{P}}_{200}(5))$ equals

$$(0.05, 0.34, 0.31, 0.24, 0.04, 0.02). \tag{2}$$

   Determine the ML estimate of p without differentiation!

3. (3p.)

   (a) Let $\mathbb{Q}_1$, ..., $\mathbb{Q}_n$ be arbitrary distributions over the finite sets $\mathcal{X}_1$, ..., $\mathcal{X}_n$, and $\mathbb{P}$ be an arbitrary distribution over $\mathcal{X}_1 \times \cdots \times \mathcal{X}_n$ with marginals $\mathbb{P}_1$, ..., $\mathbb{P}_n$. Prove that

$$D(\mathbb{P}||\mathbb{Q}_1 \times \cdots \times \mathbb{Q}_n) = D(\mathbb{P}||\mathbb{P}_1 \times \cdots \times \mathbb{P}_n) + \sum_{i=1}^{n} D(\mathbb{P}_i||\mathbb{Q}_i). \tag{3}$$

   Conclude that among the distributions $\mathbb{P}$ with marginals $\mathbb{P}_1$, ..., $\mathbb{P}_n$ the I-divergence $D(\mathbb{P}||\mathbb{Q}_1 \times \cdots \times \mathbb{Q}_n)$ is minimal if $\mathbb{P} = \mathbb{P}_1 \times \cdots \times \mathbb{P}_n$!

   (b) Let $X_1, \ldots, X_n$ be iid random variables over the set $\mathcal{X}$, and let $A \subset \mathcal{X}^n$ be an arbitrary measurable set. Prove that

$$\log Prob\,((X_1, \ldots, X_n) \in A) \le -\sum_{i=1}^{n} D(\mathbb{P}_i||\mathbb{Q}) \tag{4}$$

   where $\mathbb{Q}$ is the common distribution of X_i-s and $\mathbb{P}_i$ is the conditional distribution of X_i under the condition $(X_1, \ldots, X_n) \in A$.
   Hint: use problem 1a of the first homework set and the result of part (a) with the following choice of $\mathbb{P}$: it is the conditional joint distribution of $X_1, \ldots, X_n$ under the condition $(X_1, \ldots, X_n) \in A$.

4. (3p.) (Application of exercise 3b)

   For binary valued i.i.d. $X_1, \ldots, X_n$ with common distribution $Q = (Q(0), Q(1)) = (1-q, q)$. Let $p \le q$. Show that

$$Pr\left(\sum_{i=1}^{n} X_i \le np\right) \le 2^{-nD(p||q)} \tag{5}$$

   where

$$D(p||q) = p \log \frac{p}{q} + (1-p) \log \frac{1-p}{1-q}. \tag{6}$$

   How is this related to Sanov's theorem?

Hint: First prove that

$$Pr\left(X_i = 1 \Big| \sum_{i=1}^{n} X_i \leq np\right) \leq p \tag{7}$$

via determining

$$Pr\left(X_i = 1 \Big| \sum_{i=1}^{n} X_i = k\right),\ 0 \leq k \leq np. \tag{8}$$