Information Theory [MA5103]
Homework 3

1. Suppose we are given six symbols $\{a, b, c, d, e, f\}$ with weight $\{\frac{1}{2}, \frac{1}{10}, \frac{1}{20}, \frac{1}{6}, \frac{3}{20}, \frac{1}{30}\}$ respectively. Find a Huffman code (binary) for them.

2. Let $\mathcal{X}$ be a finite index set, $\mathcal{A}$ another finite set, and $(l_x)_{x\in\mathcal{X}}$ a sequence of nonnegative integers satisfying the Kraft inequality:
$$\sum_{x\in\mathcal{X}} |\mathcal{A}|^{-l_x} \leq 1$$
Show that there exists a prefix-free code $C : \mathcal{X} \to \mathcal{A}^+$ with the length of $C(x)$ being l_x. (Hint: construct explicitly a code tree.)

3. Let $\mathcal{X}$ be a finite set and let $\{X_1, \dots, X_n\}$ be i.i.d. random variables which take values in $\mathcal{X}$ with probability $p(x) > 0$ for $x \in \mathcal{X}$ and give random sequences $x^{(n)} = (x_1, \dots, x_n) \in \mathcal{X}^n$. First, we define the type (empirical measure) for $x^{(n)} \in \mathcal{X}^n$ by
$$t_{x^{(n)}}(x) = \frac{1}{n} N(x|x^{(n)})$$
where $N(x|x^{(n)})$ is the occurrence of x in $x^{(n)}$:
$$N(x|x^{(n)}) = |\{i \in \{1, \dots, n\} | x_i = x\}|$$
Next, for $\delta > 0$ we define the δ-strongly typical set $T_\delta^{(n)}$:
$$T_\delta^{(n)} = \left\{ x^{(n)} \in \mathcal{X}^n : \left| \frac{1}{n} N(x|x^{(n)}) - p(x) \right| \leq \delta \quad \forall x \in \mathcal{X} \right\}$$
   (a) Prove that the number of (different) types is upper-bounded by $(n+1)^{|\mathcal{X}|}$. (As you see, different $x^{(n)}$'s can have the same type.)
   (b) Prove that for $\epsilon > 0$
   $$\Pr\{X^{(n)} \in T_\delta^{(n)}\} \geq 1 - \epsilon$$
   if we take large enough $n \in \mathbb{N}$. Hint: Why not apply the law of large numbers to indicator random variables $I(X = x)$.
   (c) Prove that there exists $c > 0$ and for $x^{(n)} \in T_\delta^{(n)}$ we have
   $$2^{-n(H(X)+c\delta)} \leq p(x^{(n)}) \leq 2^{-n(H(X)-c\delta)}$$
   if we take large enough $n \in \mathbb{N}$. Here, $p(x^{(n)})$ is the probability of having the sequence $x^{(n)}$ based on the probability density $\{p(x)\}_{x\in\mathcal{X}}$ defined above.

(d) Prove that there exist $c > 0$ and for $\epsilon > 0$ we have

$$(1-\epsilon)2^{n(H(X)-c\delta)} \leq |T_{\delta}^{(n)}| \leq 2^{n(H(X)+c\delta)}$$

if we take large enough $n \in \mathbb{N}$.

4. In the setting of problem 3, we define the type class:

$$T_{\tilde{P}}^{(n)} = \{x^{(n)} \in \mathcal{X}^n : t_{x^{(n)}}(x) = \tilde{p}(x)\}$$

Here, the density $\tilde{P}$ is chosen so that $T_{\tilde{P}}^{(n)}$ is non-empty.

(a) Prove that for $x^{(n)} \in T_{\tilde{P}}^{(n)}$

$$p(x^{(n)}) = 2^{-n(H(\tilde{P})+D(\tilde{P}\|P))}$$

(b) Prove that

$$\Pr\{X^{(n)} \in T_{\tilde{P}}^{(n)}\} \leq 2^{-nD(\tilde{P}\|P)}$$