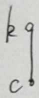
20201024 Homework

上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

1. (1) Prove $H_D(X)=\sum_{k\in L} q_k h_k$

Proof. ① If there is only one internal node, the lemma is trivially true.

② Assume it's true for all code trees with n internal nodes. When there are $n+1$ internal nodes, let k be an internal node such that k is the parent of a leaf c with maximum order. If the outcome of X is not a leaf descending from k, we identify the outcome exactly, otherwise call this random variable V ~~to be a~~ child of k. If we don't identify the outcome that isn't a leaf node of k, which happens with q_k, we further identify which of child of k is the outcome. We call this W. $X=(V,W)$. The outcome of V is a situation "n", then:



$H(V)=\sum_{k'\in L\setminus\{k\}} q_{k'}h_{k'}$. So $H(X)=H(V)+H(W|V)=\sum_{k'\in L\setminus\{k\}} q_{k'}h_{k'}+(1-q_k)\cdot 0+q_k h_k=\sum_{k\in L} q_k h_k$ $\therefore H_D(X)=\sum_{k\in L} q_k h_k$

(2) prove $L=\sum_{k\in L} q_k$.

Proof. Define $a_{ki}=\begin{cases}1 & \text{if leaf } c_i \text{ is a descendent of internal node } k.\\ 0 & \text{otherwise.}\end{cases}$

$l_i=\sum_{k\in L} a_{ki}$, $q_k=\sum_i p_i a_{ki}$

$\therefore L=\sum_i p_i l_i=\sum_i p_i \sum_{k\in L} a_{ki}=\sum_{k\in L} q_k$

2. ~~proof~~ prove the equivalence of 4 definitions of convergence in probability.

(1) $\forall \varepsilon>0, \lim\limits_{n\to\infty} P(|X_n-Y|<\varepsilon)=1$. (2) $\forall \varepsilon>0, \lim\limits_{n\to\infty} P(|X_n-Y|<\varepsilon)>1-\varepsilon$.

(3) $\forall \varepsilon>0, \lim\limits_{n\to\infty} P(|X_n-Y|\le\varepsilon)=1$. (4) $\forall \varepsilon>0, \lim\limits_{n\to\infty} P(|X_n-Y|\le\varepsilon)>1-\varepsilon$.

(1) $\Rightarrow$ (2): $\forall \varepsilon>0, \lim\limits_{n\to\infty} P(|X_n-Y|<\varepsilon)=1>1-\varepsilon$.

(2) $\Rightarrow$ (4): $\forall \varepsilon>0, \lim\limits_{n\to\infty} P(|X_n-Y|\le\varepsilon)\ge\lim\limits_{n\to\infty} P(|X_n-Y|<\varepsilon)>1-\varepsilon$.

(1) $\Rightarrow$ (3): $\forall \varepsilon>0, 1=\lim\limits_{n\to\infty} P(|X_n-Y|<\varepsilon)\le\lim\limits_{n\to\infty} P(|X_n-Y|\le\varepsilon)\le 1$

$\therefore \lim\limits_{n\to\infty} P(|X_n-Y|\le\varepsilon)=1$

(3) $\Rightarrow$ (4): $\forall \varepsilon>0, 1=\lim\limits_{n\to\infty} P(|X_n-Y|\le\varepsilon)>1-\varepsilon$.

(4) $\Rightarrow$ (1): $\forall\, 0<\delta<\varepsilon, \lim\limits_{n\to\infty} P(|X_n-Y|<\varepsilon)\ge\lim\limits_{n\to\infty} P(|X_n-Y|\le\delta)\overset{(4)}{>}1-\delta$.

$\therefore \underset{1\ge}{\lim\limits_{n\to\infty}} P(|X_n-Y|<\varepsilon)\ge\lim\limits_{\delta\to 0}\lim\limits_{n\to\infty} P(|X_n-Y|\le\delta)>\lim\limits_{\delta\to 0} 1-\delta=1$.

$\therefore \lim\limits_{n\to\infty} P(|X_n-Y|<\varepsilon)=1$.

So the four definitions are equivalent.

3. Show examples such that (1) $x\in T^n_{[X]\varepsilon}$ and $x\in W^n_{[X]\varepsilon}$.
(2) $x\notin T^n_{[X]\varepsilon}$ and $x\in W^n_{[X]\varepsilon}$. (3) $x\in T^n_{[X]\varepsilon}$ and $x\notin W^n_{[X]\varepsilon}$. (4) $x\notin T^n_{[X]\varepsilon}$ and $x\notin W^n_{[X]\varepsilon}$.

(1)

X	0	1	2
p	0.5	0.25	0.25

$x_1=(1,0,2,0)$. $H(X)=-\frac{1}{2}\log\frac{1}{2}-\frac{1}{4}\log\frac{1}{4}-\frac{1}{4}\log\frac{1}{4}=\frac{3}{2}$.

(2) $x_2=(0,1,0,1)$ $\frac{N(1;x)}{n}=\frac{2}{4}\ne p(1)$ so $x_2\notin T^n_{[X]\varepsilon}$

~~(3) if $x\in T^n_{[X]\varepsilon}$ then $x\in W^n_{[X]\varepsilon}$ because $2^{-n(H(x)+\eta)}\le p(x)\le 2^{-n(H(x)-\eta)}$~~

~~then $H(x)-\eta\le-\frac{1}{n}\log p(x)\le H(x)+\eta$, $\eta\to 0$ when $\varepsilon\to 0$ and $\eta=\varepsilon$.~~

~~$\therefore x\in W^n_{[X]\eta}$, $x\in W^n_{[X]\varepsilon}$.~~

(4) $x_4=(0,0,0)$, $\frac{N(0;x)}{n}=1\ne p(0)$ so $x_4\notin T^n_{[X]\varepsilon}$.

$-\frac{1}{n}\log p(x_4)=-\frac{1}{3}\log(0.5\times0.5\times0.5)=-\log 0.5=1\ne H(x)=\frac{3}{2}$. so $x_4\notin W^n_{[X]\varepsilon}$.

(3) let

X	0	1
p	0.1	0.9

$H(X)=-0.1\log 0.1-0.9\log 0.9\approx 0.467$

$n=100$, ~~$\delta=1$~~ $\varepsilon=0.1$

For $T^n_{[X]\varepsilon}$: $|N(0;x)-np(0)|+|N(1;x)-np(1)|\le n\varepsilon\Rightarrow 85\le N(1;x)\le 95$.

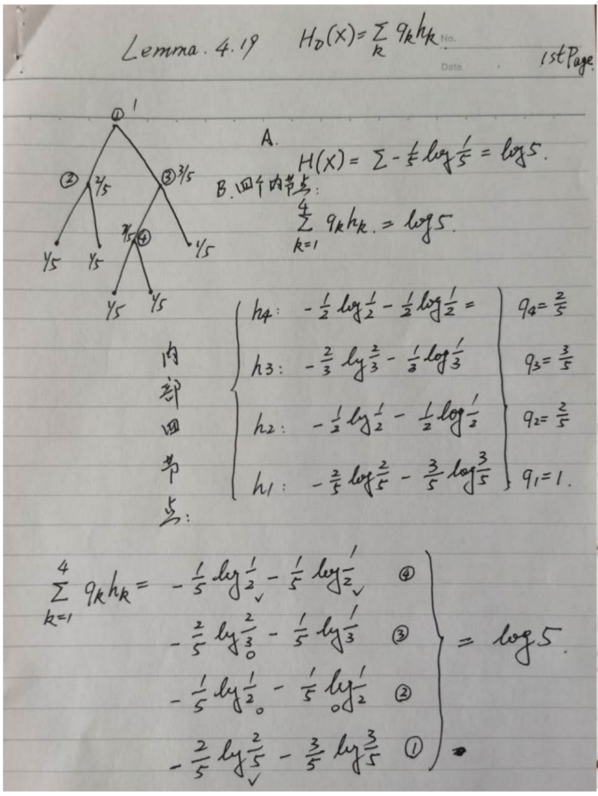
For $W^n_{[X]\varepsilon}$: $|-N(0;x)\log p(0)-N(1;x)\log p(1)-nH(x)|\le n\varepsilon\Rightarrow 87\le N(1;x)\le 93$.

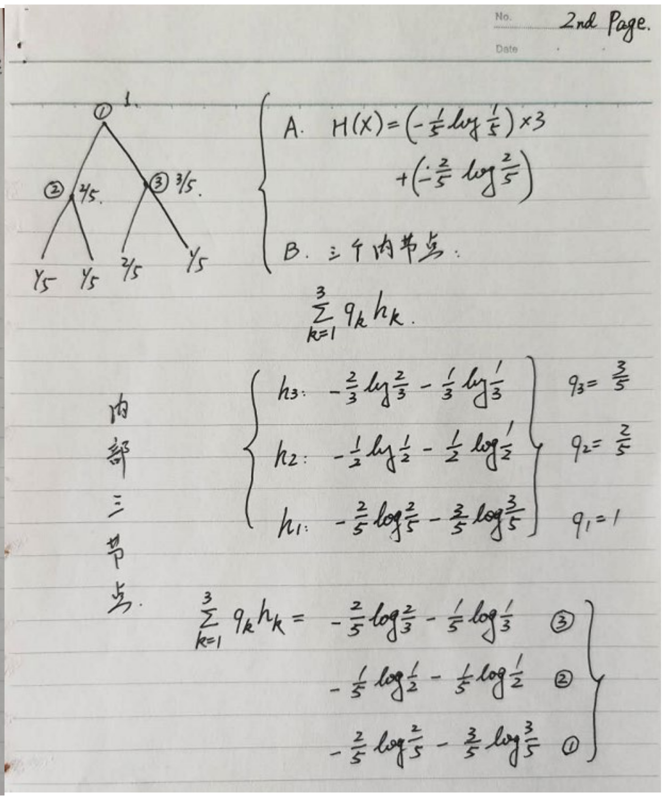
$\therefore$ When $N(1;x)=94$ or 95, $x\in T^n_{[X]\varepsilon}$ and $x\notin W^n_{[X]\varepsilon}$.

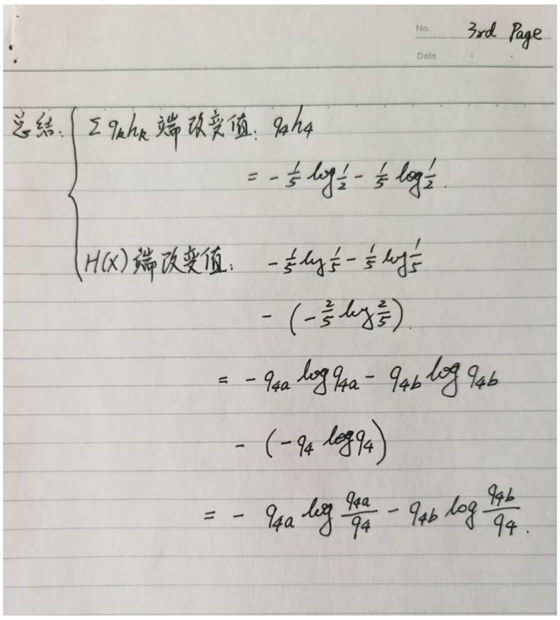
for example: $x=(0,0,0,0,0,\underbrace{1,1,\cdots,1}_{95个1})$.

1. **Prove the following two lemmas of Local Redundancy Theorem**

   ✱ **Lemma 4.19.** $\boldsymbol{H_D(X) = \sum_{k\in\mathcal{L}} q_k h_k}$. (maybe following example will show you an easier description than the textbook in the proof.)







   ✱ **Lemma 4.20.** $\boldsymbol{L = \sum_{k\in\mathcal{L}} q_k}$.

2. **Prove the equivalence of four definitions of convergence in probability.**

3. **Show examples such that**

   1) $\boldsymbol{x} \in T^n_{[X]\varepsilon}$ and $\boldsymbol{x} \in W^n_{[X]\varepsilon}$

2) $\boldsymbol{x} \notin T_{[X]\varepsilon}^{n}$ and $\boldsymbol{x} \in W_{[X]\varepsilon}^{n}$

3) $\boldsymbol{x} \in T_{[X]\varepsilon}^{n}$ and $\boldsymbol{x} \notin W_{[X]\varepsilon}^{n}$

4) $\boldsymbol{x} \notin T_{[X]\varepsilon}^{n}$ and $\boldsymbol{x} \notin W_{[X]\varepsilon}^{n}$