Homework3 郑祖彬 12112328

2.15 2.16 2.32 2.34

2.15 $I(X_1; X_2, \cdots X_n) = E_{x_1, x_2 \cdots x_n} \log \frac{p(X_1, X_2 \cdots X_n)}{p(X_1)p(X_2, \cdots X_n)} = E_{x_1, x_2 \cdots x_n} \log \frac{p(X_1)p(X_2|X_1)\cdots p(X_n|X_{n-1})}{p(X_1)p(X_2)p(X_3|X_2)\cdots p(X_n|X_{n-1})}$

$= E_{x_1, x_2 \cdots x_n} \log \frac{p(X_2|X_1)}{p(X_2)} = E_{x_1, x_2 \cdots x_n} \log \frac{p(X_1, X_2)}{p(X_1)p(X_2)} = I(X_1; X_2)$

2.16. (a) proof. since $X_1 \to X_2 \to X_3$

$I(X_1; X_3) \leq I(X_1; X_2) \leq H(X_2) \leq \log k$

since $I(X_1; X_2) = H(X_2) - H(X_2|X_1)$ and $H(X_2|X_1) \geq 0$

since $X_2 \in \{1, 2, \cdots k\}$ and unform distribution is the upper bound of $H(X_2)$.

(b) when $k=1$ $I(X_1; X_3) \leq \log k = 0$

since $I(X_1; X_3) \geq 0$

we have $I(X_1; X_3) = 0 \to X_1, X_3$ are not dependent.

2.32 (a) $\hat{X}(y) = \begin{cases} 1, & y=a \\ 2, & y=b \\ 3, & y=c \end{cases}$

~~$P_e = (1-P(X=1,Y=a)) + (1-P(X=2,Y=b)) + (1-P(X=3,Y=c))$~~

~~$= (1-\frac{1}{6}) + (1-\frac{1}{6}) + (1-\frac{1}{6}) =$~~

$P_e = (P(X=1) - P(X=1,Y=a)) + (P(X=2) - P(X=2,Y=b)) + (P(X=3) - P(X=3,Y=c))$

$= (\frac{1}{3}-\frac{1}{6}) + (\frac{1}{3}-\frac{1}{6}) + (\frac{1}{3}-\frac{1}{6}) = \frac{1}{2}$

(b) $P_e \geq \frac{H(X|Y)-1}{\log(|X|-1)} = \frac{H(X,Y)-H(Y)-1}{\log(3-1)} = \frac{\frac{3}{2}+\log 3 - \log 3 - 1}{\log 2}$

$= \frac{\frac{3}{2}-1}{1} = \frac{1}{2}$ which is the same in (a)

where $H(X,Y) = \frac{1}{6}\log 6 \times 3 + \frac{1}{12}\log 12 \times 6$ $\quad H(Y) = (\frac{1}{3}\log 3) \cdot 3$

$= \frac{1}{2}\log 6 + \frac{1}{2}\log 12$ $\quad = \log 3$

$= \frac{1}{2}(1+\log 3) + \frac{1}{2}(2+\log 3)$

$= \frac{3}{2} + \log 3$

$p(X_0, X_1 \dots X_n) = p(X_0)p(X_1|X_0)\dots p(X_n|X_{n-1})$

2.34 proof: suppose we have Markov chain $X_0 \to X_1 \to X_2 \cdots X_n$

~~$H(X_0|X_n) = H(X_0, X_n) - H(X_n)$~~

~~$H(X_0|X_{n-1}) = H(X_0, X_{n-1}) - H(X_{n-1})$~~

~~$H(X_0|X_n) - H(X_0|X_{n-1}) = H(X_0,X_n) - H(X_0,X_{n-1}) + H(X_{n-1}) - H(X_n)$~~

~~$= E_{X_0,X_n} \log \frac{1}{p(X_0,X_n)}$~~

by Data-processing inequality:

$I(X_0;X_{n-1}) \geq I(X_0;X_n)$

$H(X_0) - H(X_0|X_{n-1}) \geq H(X_0) - H(X_0|X_n)$ with n

we have $H(X_0|X_n) \geq H(X_0|X_{n-1}) \to H(X_0|X_n)$ is non-decreasing