CS258 Information Theory Homework 3.1

Zhou Litao 518030910407 F1803016

March 17, 2020

Exercise 1 (Run-length coding) Let $X_1, X_2, \ldots, X_n$ be (possibly dependent) binary random variables. Suppose that one calculates the run lengths $\mathbf{R} = (R_1, R_2, \ldots)$ of this sequence (in order as they occur). For example, the sequence $\mathbf{X} = 0001100100$ yields run lengths $\mathbf{R} = (3, 2, 2, 1, 2)$. Compare $H(X_1, X_2, \ldots, X_n)$, $H(\mathbf{R})$ and $H(X_n, \mathbf{R})$. Show all equalities and inequalities, and bound all the differences..

Solution. When $X_1, X_2, \ldots, X_n$ is determined, their running length is determined. $H(\mathbf{R}|X_1, X_2, \ldots, X_n) = 0$, which implie that

$$H(\mathbf{R}, X_1, X_2, \ldots, X_n) = H(X_1, X_2, \ldots, X_n)$$

When one element X_i is determined, given the running length, the whole sequence will be determined. That is to say $H(X_1, X_2, \ldots, X_n|X_i, \mathbf{R}) = 0$, which implies that

$$H(\mathbf{R}, X_1, X_2, \ldots, X_i \ldots, X_n, X_i) = H(\mathbf{R}, X_1, X_2, \ldots, X_n) = H(X_i, \mathbf{R})$$

Hence we have

$$\begin{aligned} H(X_1, X_2, \ldots, X_n) &= H(X_i, \mathbf{R}) && (\star) \\ &= H(\mathbf{R}) + H(X_i|\mathbf{R}) \\ &\leq H(\mathbf{R}) + H(X_i) \\ &\leq H(\mathbf{R}) + \log 2 = H(\mathbf{R}) + 1 && (\star) \end{aligned} \tag{1}$$

On the other hand, since $H(X_i|R) \geq 0$, we have

$$H(X_1, X_2, \ldots, X_n) = H(X_i, \mathbf{R}) = H(\mathbf{R}) + H(X_i|\mathbf{R}) \geq H(\mathbf{R}) \quad (\star) \tag{2}$$

The starred lines make up all the equalities and inequalities required by the problem. ☐

Exercise 2 (Grouping rule for entropy) Let $\mathbf{p} = (p_1, p_2, \ldots, p_m)$ be a probability distribution on m elements (i.e., $p_i \geq 0$ and $\sum_{i=1}^m p_i = 1$). Define a new distribution q on $m-1$ elements as $q_1 = p_1, q_2 = p_2 \cdots, q_{m-2} = p_{m-2}$, and $q_{m-1} = p_{m-1} + p_m$ [i.e., the distribution $\mathbf{q}$ is the same as $\mathbf{p}$ on $\{1, 2, \ldots, m-2\}$, and the probability of the last element in $\mathbf{q}$ is the sum of the last two probabilities of $\mathbf{p}$] Show that

$$H(\mathbf{p}) = H(\mathbf{q}) + (p_{m-1} + p_m) H\left(\frac{p_{m-1}}{p_{m-1} + p_m}, \frac{p_m}{p_{m-1} + p_m}\right) \tag{3}$$

.

Proof. By unfolding the definiton of entropy we have

$$\begin{aligned} H(\mathbf{p}) &= -\sum_{i=1}^{m} p_i \log p_i = -\sum_{i=1}^{m-2} p_i \log p_i - p_{m-1} \log p_{m-1} - p_m \log p_m \\ &= -\sum_{i=1}^{m-2} q_i \log q_i - q_{m-1} \log q_{m-1} + q_{m-1} \log q_{m-1} - p_{m-1} \log p_{m-1} - p_m \log p_m \\ &= H(\mathbf{q}) + (p_{m-1} + p_m) \log(p_{m-1} + p_m) - p_{m-1} \log p_{m-1} - p_m \log p_m \\ &= H(\mathbf{q}) + (p_{m-1} + p_m) \left(-\frac{p_{m-1}}{p_{m-1} + p_m} \log \frac{p_{m-1}}{p_{m-1} + p_m} - \frac{p_m}{p_{m-1} + p_m} \log \frac{p_m}{p_{m-1} + p_m}\right) \\ &= H(\mathbf{q}) + (p_{m-1} + p_m) H\left(\frac{p_{m-1}}{p_{m-1} + p_m}, \frac{p_m}{p_{m-1} + p_m}\right) \end{aligned} \tag{4}$$

□

Exercise 3 (Fano) We are given the following joint distribution on (X, Y) :

X \ Y	a	b	c
1	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{12}$
2	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$
3	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{6}$

Let $\hat{X}(Y)$ be an estimator for X (based on Y) and let $P_e = \Pr\{\hat{X}(Y) \neq X\}$

1. Find the minimum probability of error estimator $\hat{X}(Y)$ and the associated P_e
2. Evaluate Fano's inequality for this problem and compare.

Solution. 1. By observation, a feasible deterministic estimator for X can be defined as

$$\hat{X}(Y) = \begin{cases} 1 & Y = a \\ 2 & Y = b \\ 3 & Y = c \end{cases} \tag{5}$$

In this case, the error probability is

$$P_e = \sum_{(x,y)\in X\star Y, x\neq y} p(x, y) = 6 \times \frac{1}{12} = \frac{1}{2}$$

2. The general Fano's inequality implies that

$$P_e \geqslant \frac{H(X|Y) - 1}{\log|\mathcal{X}|} \tag{6}$$

We can calculate the conditional entropy

$$\begin{aligned} H(X|Y) &= \sum_y p(y)H(X|Y = y) \\ &= 3 \cdot \frac{1}{3} H(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}) \\ &= \frac{1}{2} + \frac{1}{2} + \frac{1}{2} = \frac{3}{2} \end{aligned} \tag{7}$$

By substituting Equation 7 into Equation 6 we know

$$P_e \geq \frac{1.5 - 1}{\log_2 3} \approx 0.3155$$

If we assume $\hat{X} : y \to x$, then by the stronger Fano's inequality we have

$$P_e \geq \frac{H(X|Y) - 1}{\log(|\mathcal{X}| - 1)} \geq \frac{1.5 - 1}{\log_2 2} = 0.5$$

Hence, the estimator we have found is the best under condition that $\hat{X} : y \to x$. It may be improved by introducing randomness. However, the P_e will not be less than 0.3155.

□

Exercise 4 (Discrete entropies) Let X and Y be two independent integervalued random variables. Let X be uniformly distributed over $\{1, 2, \ldots, 8\}$, and let $\Pr\{Y = k\} = 2^{-k}, k = 1, 2, 3, \ldots$

1. Find $H(X)$.
2. Find $H(Y)$.
3. Find $H(X + Y, X - Y)$

Solution.

1. For uniform distribution of X, $H(X) = \log|\mathcal{X}| = \log 8 = 3$
2. By definiton $H(Y) = \sum_{k=1}^{\infty} 2^{-k} \log 2^k = \sum_{k=1}^{\infty} k2^{-k} = 2$.
3. Since $(X, Y) \Leftrightarrow (X + Y, X - Y)$, we have $H(X + Y, X - Y|X, Y) = 0$ and $H(X, Y|X + Y, X - Y) = 0$. It follows that
$$\begin{aligned} H(X, Y) &= H(X + Y, X - Y|X, Y) + H(X, Y) \\ &= H(X + Y, X - Y, X, Y) \\ &= H(X + Y, X - Y) + H(X, Y|X + Y, X - Y) \\ &= H(X + Y, X - Y) \end{aligned} \tag{8}$$
Since X and Y are independent,
$$H(X + Y, X - Y) = H(X, Y) = H(X) + H(Y) = 3 + 2 = 5$$

$\square$