Indian Institute of Technology Kharagpur
Information and Coding Theory (MA41024/MA60020)
Assignment - 1, Spring Semester 2019-20

Instructor : Bibhas Adhikari

17.01.2020

1. The probability mass function of binomial distribution is given by $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$. Calculate the entropy of the random variable X.

2. A fair coin is flipped until the first head occurs. Let X denote the number of flips required. Find the entropy $H(X)$ in bits.

3. Let X_1 and X_2 be discrete random variables drawn according to probability mass functions $p_1(\bullet)$ and $p_2(\bullet)$ over the respective alphabets $X_1 = \{1, 2, \dots, m\}$ and $X_2 = \{m+1, m+2, \dots, n\}$. Let

$$X = \begin{cases} X_1 & \text{with probability } \alpha, \\ X_2 & \text{with probability } 1-\alpha. \end{cases}$$

   (a) Find $H(X)$ in terms of $H(X_1)$, $H(X_2)$, and α.

   (b) Maximize over α to show that $2^{H(X)} \leq 2^{H(X_1)} + 2^{H(X_2)}$ and interpret using the notion that $2^{H(X)}$ is the effective alphabet size.

4. Let X and Y be random variables with alphabets $X = Y = \{1, 2, 3, 4, 5\}$ and joint distribution $P(X, Y)$ is given by

$$\frac{1}{25}\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 2 & 1 & 2 & 0 & 0 \\ 2 & 0 & 1 & 1 & 1 \\ 0 & 3 & 0 & 2 & 0 \\ 0 & 0 & 1 & 1 & 3 \end{bmatrix}.$$

   Determine $H(X), H(Y), H(X|Y), H(Y|X)$, and $I(X;Y)$.

5. What is the minimum value of $H(p_1, \dots, p_n) = H(P)$ as P ranges over the set of n-dimensional probability vectors? Find all P that achieve this minimum.

6. Let $C_\alpha = \sum_{n=2}^{\infty} \frac{1}{n(\log(n))^\alpha}$. Prove that

   (a) $C_\alpha \leq \infty$ if $\alpha > 1$ and , $C_\alpha = \infty$ if $0 \leq \alpha \leq 1$.

   (b) $p_\alpha(n) = [C_\alpha n(\log(n))^\alpha]^{-1}$ for $n = 2, 3, \dots \infty$ is a probability distribution.

   (c) $H(p_\alpha) < \infty$ for $\alpha > 2$ and $H(p_\alpha) = \infty$ for $1 < \alpha \leq 2$.

7. Prove that if f is a twice differentiable convex function in an interval I, then $f''(x) \geq 0$. Using this criterion verify which of these functions $f(x) : (0, \infty) \to \mathbb{R}$ given by $f(x) = log(x), xlog(x)$ are convex or concave.

8. Let $a_1, a_2, \dots a_n$ and $b_1, b_2, \dots b_n$ be two sets of positive numbers. Prove that

$$\sum_{i=1}^{n} a_i \log\left(\frac{a_i}{b_i}\right) \geq \left[\sum_{i=1}^{n} a_i\right] \log\left(\frac{\sum_{i=1}^{n} a_i}{\sum_{i=1}^{n} b_i}\right).$$

   Equality occurs when $\frac{a_i}{b_i} = \text{constant}$.

9. Prove that $H(p)$ is concave in p, that is, $\lambda H(p_1) + (1-\lambda)H(p_2) \leq H(\lambda p_1 + (1-\lambda)p_2)$ holds for $0 \leq \lambda \leq 1$.

10. Let X be a random variable taking on a finite number of values. What is the (general) inequality relationship of $H(X)$ and $H(Y)$ if $Y = 2^X$, and $Y = \cos(X)$.

11. Let X be a discrete random variable. Show that the entropy of a function $g(X)$ is less than or equal to the entropy of X, that is, $H(g(X)) \leq H(X)$.

12. Under what conditions $H(X|g(Y)) = H(X|Y)$?

13. Show that if $H(Y|X) = 0$, then Y is a function of X.

14. Prove that for $n \geq 2$, $H(X_1, X_2, \ldots, X_n) \geq \sum_{i=1}^{n} H(X_i|X_j; i \neq j)$.

15. Prove that $\frac{1}{2}[H(X_1, X_2) + H(X_2, X_3) + H(X_1, X_3)] \geq H(X_1, X_2, X_3)$.

16. Show that the entropy of the probability distribution, $(p_1, \ldots, p_i, \ldots, p_j, \ldots, p_m)$, is less than the entropy of the distribution $(p_1, \ldots \frac{p_i+p_j}{2}, \ldots \frac{p_i+p_j}{2}, \ldots, p_m)$.

17. Let X and Y be random variables that take on values $x_1, x_2, \ldots, x_r$ and $y_1, y_2, \ldots, y_s$, respectively. Let $Z = X + Y$.
    (a) Show that $H(Z|X) = H(Y|X)$. Argue that if X, Y are independent, then $H(Y) \leq H(Z)$ and $H(X) \leq H(Z)$. Thus, the addition of independent random variables adds uncertainty.
    (b) Give an example of (necessarily dependent) random variables in which $H(X) > H(Z)$ and $H(Y) > H(Z)$.
    (c) Under what conditions does $H(Z) = H(X) + H(Y)$?

18. Prove that the function $\rho(X,Y) = H(X|Y) + H(Y|X)$ is a metric on the set of all random variables defined on same sample space.

19. For the metric $\rho(X,Y)$ given in the above question prove that

$$\begin{aligned} \rho(X,Y) &= H(X) + H(Y) - 2I(X;Y) \\ &= H(X,Y) - I(X;Y) \\ &= 2H(X,Y) - H(X) - H(Y). \end{aligned}$$

20. Do $I(X;Y) = 0$ and $I(X;Y|Z) = 0$ imply each other? If so, give a proof. If not , give a counterexample.

21. Give an example for which $D(\bullet||\bullet)$ does not satisfy the triangular inequality.

22. Prove that $D(p||q)$ is convex in the pair (p,q), that is, if (p_1, q_1) and (p_2, q_2) are two pairs of probability distributions on a common alphabet, then

$$D(\lambda p_1 + (1-\lambda)p_2||\lambda q_1 + (1-\lambda)q_2) \leq \lambda D(p_1||q_1) + (1-\lambda)D(p_2||q_2).$$

23. Let $p(x,y)$ and $q(x,y)$ be two probability mass functions on $X \times Y$. Prove that $D(p(x,y)||q(x,y)) \geq D(p_x||q_x)$.