CIE 6020 Assignment 1

WU, Chenhao 117010285

January 25, 2019

1. If the base of the logarithm is b, we denote the entropy as $H_b(X)$. Show that $H_b(X) = (\log_b a)H_a(X)$.

Proof:

$$\begin{aligned}
(\log_b a)H_a(X) &= (\log_b a)\sum_{x\in\mathcal{X}} p(x)\log_a p(x) \\
&= \sum_{x\in\mathcal{X}} p(x)(\log_b a)\log_a p(x) \\
&= \sum_{x\in\mathcal{X}} p(x)(\log_b a^{\log_a p(x)}) \\
&= \sum_{x\in\mathcal{X}} p(x)\log_b p(x) \\
&= H_b(X)
\end{aligned}$$

2. *Coin flips.* A fair coin is flipper until the first head occurs. Let X denote the number of flips required.

(a) Find the entropy $H(X)$ in bits. The following expressions may be useful:

$$\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}$$

$$\sum_{n=0}^{\infty} nr^n = \frac{r}{(1-r)^2}$$

(b) A random variable X is drawn according to this distribution. Find an "efficient" sequence of yes-no questions of the form, "Is X contained in the set S?" Compare $H(X)$ to the expected number of questions required to determine X.

Answer:

(a): The probability mass function of X: $p_X(n) = P(X = n) = (\frac{1}{2})^{n-1}\frac{1}{2} = (\frac{1}{2})^n$

$$\begin{aligned} H(X) &= -\sum_{i=1}^{\infty}(\frac{1}{2})^i \log(\frac{1}{2}^i) \\ &= -\sum_{i=1}^{\infty}(\frac{1}{2})^i i \log(\frac{1}{2}) \\ &= \sum_{i=1}^{\infty} i(\frac{1}{2})^i \\ &= 2 \end{aligned}$$

(b): Since the pmf of X is exponentially decreasing, one of the reasonable questions for nth question is "Is X = n?". Let Y denote the number of questions need to ask to determine the exact number of flips, then the probability mass function of Y can be given by

$$p_Y(n) = P(X = n | X \geq n) = (1 - \sum_{i=1}^{n-1} p(x))(\frac{1}{2})^n = (\frac{1}{2})^n$$

and therefore, the expectation of Y can by given by

$$\begin{aligned}
E[Y] &= \sum_{i=1}^{\infty} i p_Y(i) \\
&= 2 \\
&= H(X)
\end{aligned}$$

From the equivalence of $E[Y]$ and $H(X)$ we can infer that this sequence of questions are optimal, since it can be proved that each nth question can get 1 bit information from the set of all possible solutions.

3. *Entropy of functions.* 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

(a) $Y = 2^X$?

(b) $Y = cos(X)$?

Answer:

(a) Suppose that x's alphabet $\mathcal{X} = (x_1, x_2, ..., x_m)$ and y's alphabet $\mathcal{Y} = (y_1, y_2, ..., y_n)$

For $Y = f(X) = 2^X$, $f : \mathcal{X} \mapsto \mathcal{Y}$ is a one-to-one mapping, and therefore by definition

$$\begin{aligned}
H(X) &= -\sum_{x \in \mathcal{X}} p(x) \log p(x) \\
&= -\sum_{y} \sum_{x : f(x) = y} p(x) \log p(x) \\
&= -\sum_{y \in \mathcal{Y}} p(y) \log p(y) \\
&= H(Y)
\end{aligned}$$

(b) Suppose that x's alphabet $\mathcal{X} = (x_1, x_2, ..., x_m)$ and y's alphabet $\mathcal{Y} = (y_1, y_2, ..., y_n)$

Intuitively, for $Y = f(X) = cos(X)$, $f : \mathcal{X} \mapsto \mathcal{Y}$ is surjective but not injective

$$
\begin{aligned}
H(X) &= -\sum_{x \in \mathcal{X}} p(x) \log p(x) \\
&= -\sum_{y} \sum_{x: f(x)=y} p(x) \log p(x) \\
&> -\sum_{y} \sum_{x: f(x)=y} p(x) \log p(y) \\
&= -\sum_{y} p(y) \log p(y) \\
&= H(Y)
\end{aligned}
$$

Therefore, $H(X) > H(Y)$ for $Y = cos(X)$

4. What is the minimum value of $H(p_1, ..., p_n) = H(\mathbf{p})$ as $\mathbf{p}$ ranges over the set of n-dimensional probability vectors? Find all $\mathbf{p}$'s that achieve this minimum

Answer: The entropy of $\mathbf{p}$ is given by

$$
\begin{aligned}
H(\mathbf{p}) &= \sum_{i=1}^{n} p_i \log p_i \\
&= \sum_{i=1}^{n} p_i \log \frac{p_i}{1} \\
&\geq (\sum_{i=1}^{n} p_i) \log \frac{\sum_{i=1}^{n} p_i}{n}
\end{aligned}
$$

Notice that $\sum_{i=1}^{n} p_i = 1$, and therefore $H(\mathbf{p}) \geq \log \frac{1}{n}$ with the equality being held iff $p_i n = 1$

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

Proof: