

# Homework #1

*Jonathan McFadden*

TCSS - 580 : Winter 2018  
Information Theory

**Problem 2.3):** Let  $\mathbb{P}^n$  be the set of all  $n$ -dimensional probability vectors, with elements  $\vec{p} \in \mathbb{P}^n$  defined as  $\vec{p} = (p_1, p_2, \dots, p_i, \dots, p_n)$  for  $i \in \mathbb{Z}^+ \ni i \leq n$ . By the definition of a probability space, we must have

$$\vec{p} \cdot \vec{1} = \sum_{i=1}^n \{p_i\} = 1 \quad , \quad \forall \vec{p} \in \mathbb{P}^n, \quad (2.3-1)$$

where vector  $\vec{1}$  is defined as  $\vec{1} = (q_1, q_2, \dots, q_k, \dots, q_n) \in \mathbb{Z}^n$  with  $q_k = 1, \forall k \in [1, n]$  (where the interval  $[1, n]$  is defined such that  $[1, n] \subseteq \mathbb{Z}^+$ ). Furthermore, the definition of a probability space also requires that, for any  $\vec{p} \in \mathbb{P}^n$ , the elements of  $\vec{p}$  (the  $p_i \in \vec{p}$  such that  $i \in \mathbb{Z}^+ \ni i \leq n$ ) satisfy the condition

$$p_i \geq 0 \quad (2.3-2)$$

for all  $i \in \mathbb{Z}^+ \ni i \leq n$ .

The expression in 2.3-1 guarantees that the  $p_i$  of any  $\vec{p} \in \mathbb{P}^n$  satisfy the bound  $0 \leq p_i \leq 1$  where  $i \in \mathbb{Z}^+ \ni i \leq n$ . Therefore, the relation

$$p_i \log_2 [p_i] \geq 0 \quad (2.3-3)$$

holds for all  $p_i$  of any  $\vec{p} \in \mathbb{P}^n$ . Moreover, for the cases where  $p_i = 0$  or  $p_i = 1$ , it is clear that the expression in 2.3-3 reduces to equality. Specifically, the relation  $p_i \log_2 [p_i]$  becomes

$$p_i \log_2 [p_i] = 0 \quad (2.3-4)$$

for the case where  $p_i = 0$  or  $p_i = 1$ . Moreover, the relation in 2.3-4 also represents the **smallest** possible value/result for the expression  $p_i \log_2 [p_i]$ . That is to say, that when  $p_i = 0$  or  $p_i = 1$ , then  $p_i \log_2 [p_i]$  is at a minimum.

The result in 2.3-1, makes it is clear that *only* **ONE**  $p_i$  in each  $\vec{p} \in \mathbb{P}^n$  may have the value  $p_i = 1$ ; therefore the probability vectors  $\vec{p} \in \mathbb{P}^n$  which result in a minimum value for  $p_i \log_2 [p_i]$  all have exactly one non-zero element

with the non-zero element having a value of one. This implies that there are only  $n$  such probability vectors,  $\vec{p}^*$  within any  $\mathbb{P}^n$ . Furthermore, the value of  $H(X) = \sum_{i=1}^n \{p_i \log_2 [p_i]\}$  for any such  $\vec{p}^*$  is also zero.

**Problem 2.4 a):** Recall the chain-rule for conditional entropies of  $X$  given  $Y$ ,

$$H(X | Y) = H(X, Y) - H(Y) \quad (2.4-1)$$

We apply the expression in 2.4-1 to the case of  $g(X)$  given  $X$  to obtain

$$H(g(X) | X) = H(g(X), X) - H(X) \quad (0.0.1)$$

by rearranging the expression in the previous result as follows,

$$H(g(X), X) = H(X) + H(g(X) | X) \quad (2.4-2)$$

we obtain the desired result.

**Problem 2.4 b):** For any given value of  $X$ , we automatically know  $g(X)$ . Therefore, the expression for  $H(g(X), X)$  in 2.4-2 becomes

$$H(g(X), X) = H(X)$$

which is the desired result.