

Note on PS2 Q3

Tom Kealy

February 17, 2015

Abstract

This is a short note on methods of proof required to answer IT problem sheet 2, question 3.

1 The problem

Definition 1.1 (Problem 3). *Show that the expected length of a Shannon code is equal to $H(P)$ if and only if there exist natural numbers, $\{n_x\} \in \mathbb{N}$, $x \in \mathcal{X}$, such that $P(x) = 2^{-n_x}$.*

2 Solution

There are several solutions that you can use to solve problems like this. One is to form a chain of equivalences between the two statements. The other is to prove the theorem one way and then prove it in the opposite direction. The catch with this way, is that you must supply different ideas for the direct and converse parts..

An example of the first way is to demonstrate the positivity of $\mathbb{E}_P(l \circ f) - H(P)$, (with $l(f(x)) = \lceil -\log P(x) \rceil$), and then using the inequalities for the ceiling function to deduce the conditions where equality occurs. This is covered in more detail in the first section.

Most students followed the second method, with ideas looking something like:

Proof. Example proof

(\leftarrow) by direct computation. Let $P(x) = 2^{-n_x}$ for some $\{n_x\} \in \mathbb{N}$. By Shannon's LSCT (direct part) \exists a uniquely decodable code (UDC) with $l_x = \lceil -\log P(x) \rceil \forall x \in \mathcal{X}$.

In this case $l_x = -\log P(x)$, $\forall x \in \mathcal{X}$, as $-\log P(x) \in \mathbb{N}$ for some $\{n_x\} \in \mathbb{N}$ by assumption (we can choose these $\{n_x\}$ to coincide with the requirements for the Shannon code). So,

$$\mathbb{E}_P(l \circ f) = \sum_{x \in \mathcal{X}} P(x) l_x = \tag{1}$$

$$= - \sum_{x \in \mathcal{X}} P(x) \log P(x) = H(P) \tag{2}$$

(\rightarrow) Let $\mathbb{E}_P(l \circ f) = H(P)$. So,

$$\mathbb{E}_P(l \circ f) = \sum_{x \in \mathcal{X}} P(x) \lceil -\log P(x) \rceil = \quad (3)$$

$$= \sum_{x \in \mathcal{X}} P(x) (-\log P(x)) = H(P) \quad (4)$$

This is equivalent to $-\log P(x) = \lceil -\log P(x) \rceil$ i.e. $-\log P(x)$ is an integer, which implies $P(x)$ is a power of 2. \square

This proof fails because while two sums may be identical, the terms needn't be identical. For example:

$$\sum_{n=1}^{\infty} \frac{1}{2^n} = \sum_{n=1}^{\infty} \frac{6}{\pi n^2} = 1 \quad (5)$$

but it's not true that

$$\frac{1}{2^n} = \frac{6}{\pi n^2} \quad (6)$$

However, this proof is only checking that the statement is true as the same idea is used in both the direct and converse part. Also, absent are any ideas about UDCs.

An example of a correct proof with a direct and converse part is:

Proof. ((\leftarrow)) by direct computation. Let $P(x) = 2^{-n_x}$ for some $\{n_x\} \in \mathbb{N}$. By Shannon's LSCT (direct part) \exists a uniquely decodable code (UDC) with $l_x = \lceil -\log P(x) \rceil \forall x \in \mathcal{X}$.

In this case $l_x = -\log P(x)$, $\forall x \in \mathcal{X}$, as $-\log P(x) \in \mathbb{N}$ for some $\{n_x\} \in \mathbb{N}$ by assumption (we can choose these $\{n_x\}$ to coincide with the requirements for the Shannon code). So,

$$\mathbb{E}_P(l \circ f) = \sum_{x \in \mathcal{X}} P(x) l_x = \quad (7)$$

$$= - \sum_{x \in \mathcal{X}} P(x) \log P(x) = H(P) \quad (8)$$

(\rightarrow) From the converse part to Shannon's LSCT we know:

$$H(P) \leq \mathbb{E}_P(l \circ f) \quad (9)$$

for any UDC. In the proof of the theorem, Kraft's inequality and the log-sum inequality are used to demonstrate this result. Therefore, equality occurs when $\sum_{x \in \mathcal{X}} 2^{-l(f(x))} = 1$ and $\exists c \geq 0 : P(x) = c 2^{-l(f(x))}$. This implies $c = 1$ from summing over both sides of the second condition. Then $P(x) = 2^{-l(f(x))}$ and because $l_x \in \mathbb{N}$ from the definition of the length function, we are done. \square