A COMBINATORIAL APPROACH TO THE SHANNON ENTROPY

Aidan Rocke
aidanrocke@gmail.com

April 7, 2021

ABSTRACT

Within the setting of information theory, we typically assume that the Shannon Entropy defines the uncertainty associated with a random variable. However, in this article I demonstrate that the Shannon Entropy may be defined in purely combinatorial terms as a state function which measures the number of possible arrangements. In doing so, we generalise its application beyond random variables and save ourselves the trouble of having to define what is a random variable.

1 The Combinatorial Entropy S_c

Let's suppose you have a small deck of cards in your hands with four kings, three aces, and two queens. If you'd like to consider the number of rearrangements of this deck you might want to know how many unique permutations are possible assuming that cards of the same type are indistinguishable from each other.

In principle, mathematicians may rely upon portfolio theory which may be considered a rational approach to gambling. Given a finite number of assets, a mathematician will construct a portfolio by taking a weighted average of these assets $W = \frac{9!}{4!3!2!}$, and in general this multinomial coefficient is a solution of the general case:

$$W = \frac{N!}{\prod_i N_i!} \tag{1}$$

where N is the total number of cards and N_i is the number of cards of a particular type.

Now, we may observe that all permutations may be enumerated and assigned a number(in binary) from 0 to W-1 so the string of $\log_2(W)$ bits may be used to encode each permutation. This allows us to define the Combinatorial Entropy S_c as the average number of bits per permutation:

$$S_c = \frac{\log_2(W)}{N} = \frac{1}{N} \log_2 \left(\frac{N!}{\prod_i N_i!} \right) \tag{2}$$

and the reader may infer that S_c is in some sense a measure of diversity as the magnitude of S_c varies inversely with the number of cards of the same type:

$$\frac{\partial S_c}{\partial N_i} = -\frac{1}{N} \cdot \frac{1}{N_i} \tag{3}$$

and what is even more interesting is that the Shannon Entropy may be derived as an approximation of S_c so the Shannon Entropy $H(\cdot)$ inherits the properties of S_c.

2   A combinatorial approach to the Shannon Entropy

Using Stirling's log-factorial approximation $\ln N! \approx N \ln N - N$, we have:

$$S_c \approx \frac{1}{N} \Big(\sum_i N_i \log_2 N - \sum_i N_i \log_2 N_i \Big) \tag{4}$$

since $\sum_i N_i = N$. Now, we find:

$$S_c \approx - \Big(\sum_i \frac{N_i}{N} \log_2 \frac{N_i}{N} \Big) \tag{5}$$

and if we define the frequencies $p_i = \frac{N_i}{N}$ we recover the usual Shannon entropy $H(\cdot)$:

$$H(X) = - \sum_i p_i \log_2 p_i \tag{6}$$

where X refers to the current state of the card deck and $P = \{p_i\}_{i=1}^n$ is a discrete probability distribution. It follows that $H(\cdot)$ has a natural interpretation as a state function that maps the current state X to a combinatorial measure of diversity.

3   Discussion

Finally, it is worth noting that the expectation operator $\mathbb{E}$ is generally defined only for random variables. However, I will note that we may define (6) using the expectation:

$$H(X) = \mathbb{E}[-\log P(X)] = - \sum_i p_i \log_2 p_i \tag{7}$$

so the expectation operator may be defined without having to define the notion of a random variable. This point is worth emphasizing because randomness is actually a subtle notion which is typically poorly defined within classical mathematics if it is defined at all.

References

[1] Edwin Jaynes. Information Theory and Statistical Mechanics I. The Physical Review. 1957.

[2] Edwin Jaynes. Information Theory and Statistical Mechanics II. The Physical Review. 1957.