Documentation

Jerry Wang

Contents

1 Introduction

1.1 Entropy - Surprise Factor

First we would like to define the concept of the "surprise" of an event E occurring and define entropy based off that. We deem this important since entropy plays a large role in many of the metric developed and implemented. As such, it would be helpful to define a consistent mathematical framework as well as develop intuition for these definitions at a precise mathematical level.

The following methods and definitions are based on Ross's A First Course in Probability [1]. Mathematically, it makes sense that the surprise invoked by an event E occurring should be a function of the probability of event E itself, which we will denote p. Thus we define $S(p)$ as the surprise invoked by an event with a probability of p. Now we will begin by stating some axioms for this definition of surprise.

Axiom 1

$$S(1) = 0$$

Intuitively, that just means we should feel no surprise when a event with probability 1 occurs. That is, it is not surprising at all for a sure event to occur.

Axiom 2

$$p < q \implies S(p) < S(q)$$

That is, $S(p)$ is a strictly decreasing function of p.

Axiom 3

$$S(p) \text{ is a continuous function with respect to } p$$

Axiom 4

$$S(pq) = S(p) + S(q)$$

The intuition of this axiom is given when we consider independent events E_1 and E_2 with probabilities of occurring with p and q respectively. Since E_1 and E_2 are independent, $P(E_1 \cap E_2) = P(E_1) \cdot P(E_2) = pq$. Thus the surprise invoked by $E_1 \cap E_2$ should be defined as $S(pq)$. But now, let us consider that we are first told that event E_1 occurred, and then afterwards event E_2 occur. We know the total surprise invoked is simply the surprised invoked by $E_1 \cap E_2$ which is $S(pq)$, and since we know that $S(p)$ is the surprise invoked by event E_1 alone, it follows that $S(pq) - S(p)$ should represent the initial surprise invoked by E_2. Due to independence, we know the probability of E_2 is still q and thus the surprise of event E_2 should remain $S(q)$. Thus we have that $S(pq) - S(p) = S(q)$ or that $S(pq) = S(p) + S(q)$.

From the axioms, we can prove that

$$S(p) = -C \log_2 p$$

where $C > 0$. From Axiom 4, we have that

$$S(p^2) = S(p) + S(p) + 2S(p)$$

From induction, we also have that for $m \in \mathbb{Z} > 0$

$$S(p^m) = mS(p) \tag{1}$$

Also note that for $n \in \mathbb{Z} > 0$

$$S(p) = S\left((p^{\frac{1}{n}})^n\right) = nS(p^{\frac{1}{n}})$$

which implies that

$$S(p^{\frac{1}{n}}) = \frac{S(p)}{n} \tag{2}$$

Combining equations 1 and 2, we can define

$$S(p^x) = xS(p) \tag{3}$$

for $x \in \mathbb{Q} > 0$. Now from Axiom 3, we can define equation 3 $\forall x \in \mathbb{Q} > 0$.

Now let us take $x = -\log_2 p$ which implies $p = \left(\frac{1}{2}\right)^x$ which allows for the following relation.

$$S(p) = S\left((\frac{1}{2})^x\right) = xS\left(\frac{1}{2}\right) = -S\left(\frac{1}{2}\right) \cdot \log_2 p$$

Now notice $S\left(\frac{1}{2}\right) = C$ is simply a constant (that is non-zero thanks to Axiom 1 and Axiom 2).

Thus we have shown that

$$S(p) = -C \log_2 p \tag{4}$$

follows from the axioms. Lastly, note that it is standard to let $C = 1$.

2 Mathematical Metrics

2.1 Shannon Entropy Metric

Now we aim to quantify the expected amount of surprise incurred by a random variable X. Since $\log_2 p_i$ is the value of the surprise invoked when an event (let's say E_i) with probability p_i occurs, the expected value of $S(p_i)$ for any outcome of E_i is $p_i \cdot \log_2 p_i$. Now summing over all expectations, we have the expected surprise of a random variable X which takes on n values with probability $p_1, p_2, ...p_n$ is

$$H(x) = \sum_{i=1}^{n} p_i \log_2 p_i$$

2.1.1 Shannon Entropy Ratio

It can be proven using Lagrange Multipliers that the distribution that maximizes entropy such that . Thus we will take a score

$$R = \frac{H(x,n)}{H_{\max}(n)}$$

2.2 Password Entropy Metric

References

[1] Ross, Sheldon M. 2019. A First Course in Probability. 10th ed. Pearson.