Information and Entropy

Chapter 5: Probability

1. Events

There's some Concepts:

- **Universal event**: $p(\text{universal event}) = 1$
- **Null event**: $p(\text{null event}) = 0$
- **Mutually exclusive**: $A \cap B = \emptyset$
- **Exhautive**: $A \cup B = S$, where S is the sample space.
- **Partition**: Both mutually exclusive and exhaustive.

2. Joint Events and Conditional Probabilities

Assume event A and B are independent, the **joint event** can be found:

$$p(A, B) = p(A)p(B)$$

We can also use conditional probabilities to describe:

$$p(A, B) = p(B)p(A \mid B) = p(A)p(B \mid A)$$

And It's also known as Bayes' Theorem.

3. Averages

The average value (expected value) H_{av} can be found as:

$$E_{av} = \sum_i p(A_i)c_i$$

Where A_i is a event, and c_i is its actual value.

4. Information

We will learn information when we know a event happened. The information learned form outcome i is:

$$\log_2\left(\frac{1}{p(A_i)}\right)$$

.

For all outcomes, we can take their average information:

$$I = \sum_i p(A_i)\log_2\left(\frac{1}{p(A_i)}\right)$$

It's called **entropy** of a source.

Basically, more likely the event probabilities are, higher value the entropy will get. It's like we playing the guess-the-number game. Assume the number is in 1-100, the best way is to guess 50, so we can learn if it's above 50 or under 50.