

Bayesian Indicator

Bayesian Indicator

Simple Bayesian Indicator

We define a large window which is of length k :

$$[P_i, P_{i+1}, \dots, P_{i+k}]$$

We define a smaller window which is of length j :

$$[P_{i+k-j}, \dots, P_{i+k}]$$

Now we create a transition probability matrix based on the smaller window.

$$P(U_t \rightarrow U_{t+1})$$

$$P(U_t \rightarrow D_{t+1})$$

$$P(D_t \rightarrow U_{t+1})$$

$$P(D_t \rightarrow D_{t+1})$$

Based on the larger window we calculate the overall probabilities:

$$P(U)$$

$$P(D)$$

We now calculate:

$$P(S_{t+1}|S_t) = \frac{P(S_t|S_{t+1})P(S_t)}{\sum_{i=0}^{i=1} P(S_t|S_{t+1}^i)P(S_{t+1}^i)}$$

Where $S = (U, D)$ and $S_0 = U$ and $S_1 = D$.

An extension to this simple indicator, to take into account the actual length of sequences is to select an even larger window and look at the frequency of subsequences and there lengths. We treat $P(S_{t+1}|S_t)$ as a hypothesis.

So:

$$P(H_i) = P(S_{t+1}^i|S_t)$$

Where S^i is the state of U, D .

$$P(H_i|e) = \frac{P(e|H_i)P(H_i)}{\sum_{i=0}^{i=1} P(e|H_i)P(H_i)}$$

Now e is calculated based on the length of the current subsequence it is on.

Now, we take a large window of size N .

There is a loss of information as essentially everything is encoded in 1's and 0's. In conjunction with this, we should be looking at returns.

Now say we have a collection of Down and Up sequence lengths, which is associated to its frequency. For example:

$$\begin{aligned}Down[i - 1] &= DownFrequency(i) \\Up[i - 1] &= UpFrequency(i)\end{aligned}$$

Whereby i , represents the sub sequence length we are looking for.

Now note that we can only have H_1 and H_2 . We have to generate the probability for both hypothesis considering the position we are in. H_1 represent given the current state, the price will increase. Meanwhile H_2 represents given the current state, the price will decrease.

Now the current state is important, the current state affects whether we will be looking at UpFrequency or DownFrequency, we identify what current state we are in and we count the length of our current states so far.

Once, we recognise what current state we are in we can now calculate:

$$P(e|H_i)$$

This is saying, given the hypothesis does occur whats the probability of the evidence? Well, if we assume say $H_i = H_1$, there is an increase in price and if we assume our current state is Up , so we are currently on an "Up Streak". If the "Up Streak" length so far is lets say 5.

Likelihood Probability : Calculation *****

C = Current State L = Length of Current State

if $C == UP$ then: Calculate Probability of the evidence as being of length $L + 1$ given C , since the hypothesis is that there is an increase in price.

$$P(e|H_1) = P(L + 1|C)$$

Calculate the Probability of the evidence as being of length L given C , since the hypothesis is that there is a decrease in price, so the streak length has ended there.

$$P(e|H_2) = P(L|C)$$

Similarly, if $C == DOWN$ then: Calculate the probability of the evidence as being of length L given C , since the hypothesis is that there is an increase in price. (Breaks streak of price decreasing)

$$P(e|H_1) = P(L|C)$$

Calculate the Probability of the evidence as being of length L given C , since the hypothesis is that there is a decrease in price, so the streak length is of the price going down is going to increase.

$$P(e|H_2) = P(L + 1|C)$$

So now that we can calculate:

$$P(H_i|e) = \frac{P(e|H_i)P(H_i)}{\sum_{i=0}^{i=1} P(e|H_i)P(H_i)}$$

We can also look at bayes factor to determine the strength of such changes:

$$\frac{P(H_1|e)}{P(H_2|e)} = \frac{P(e|H_1)P(H_1)}{P(e|H_2)P(H_2)}$$

So if say $B_{12} = 1.49$ then the evidence points towards choosing UP instead.

The main purpose of this construction is to sort of take into consideration the current streaks in price directions as historically there is a record of what the frequency of each streak length would be. If the simple bayesian indicator was simply presented in such a way that the probability of the price going up is far greater than the probability of the price going down, then it would always in most cases indicate that the price is going to go up. However, if you keep the length of streaks in mind you can sort of reduce the effect the two windows have based on there small sample of information. The length of streaks is an example of how a lot of information can also be effectively condensed and used.

However, both these indicator focus on ONLY price directions, there is a lot of information which is missing such as low, high, close, open, general volatility and volume.

So how can we incorporate this sort of system while using actual numerical values?

Instead of encoding changes in price by 0's and 1's we can encode it using log returns. We can still use streaks but along with it we can also use historical data which looks at the log returns from the beginning of the streak to the end of the streak too. We can look at the historical distribution of log returns for each differing lengths of streaks too.

More will be written about this method later on.