Approach #1 - Testing if Data Fits a Markov Chain

Question

Can the day to day price fluctuations of bitcoin, UP and DOWN, be considered (meaningfully) as the output of a markov chain?

Background Research

My background research consisted of downloading the day to day price history of bitcoin from http://www.coindesk.com/price/.

Hypothesis

I hypothesize that the UP and DOWN day to day price fluctuations of bitcoin cannot be expressed (meaningfully) as the output of a markov chain.

Procedure

Step 1 - Clean the Data

The first task is to clean the data. I believe that only the price history from 2013 and later should be considered. The price history prior to 2013 should be discarded for the following reasons: 1) It was early in the cryptocurrency's history and I do not have confidence that the record kept were suitably accurate. 2) The dataset will be plenty large without pre-2013 data, and the absence of pre-2013 data will not hinder testing the hypothesis, therefore there is no reason not to err on the side of caution and exclude pre-2013 data.

Step 2 - Assign Classes to the Data

The second task is to assign classes to the data. The class on day i will be assigned “UP” if the following inequality is true. Otherwise the class on day i will be assigned “DOWN”:

Class(i) == “UP” if Price(i+1) > d * Price(i)

Here, d will be a value from the set {1.00, 1.01, …, 1.05}. There will exist a dataset for each value of d, and classes will be assigned to each dataset as described above. Therefore, there will ultimately be six datasets in this experiment.

The purpose of the varying d values is to act as a buffer to slight changes in price. For example, when d=1.00, a price change of $100.00 to $100.01 would technically be labeled "UP". However, a price change from $100.00 t0 $100.00 would be labeled "DOWN. The d value is increased from dataset to dataset to provide a buffer against noise which might cause slight price changes from day to day.

Step 3 - Prepare the Data for Analysis

A markov chain consists of states and transition properties. For this experiment, the basic structure of the markov model we are considering will look like the image below:

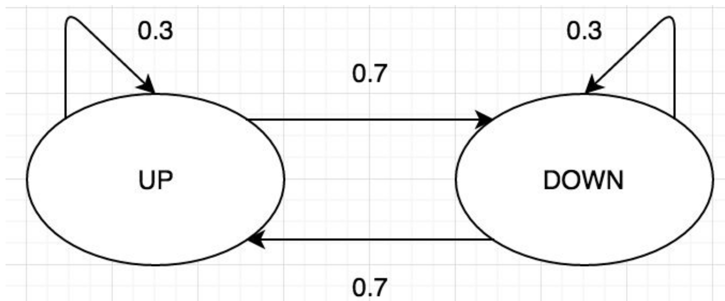

Figure 1: Basic Markov Chain

In the above image, if the state on day i is UP, then the probability of the state being UP on day i+1 is 30%, and the probability of the state being DOWN on day i+1 is 70%.

Perform this step on all six datasets.

The third task is to iterate through the dataset and make a list of all transitions from the UP state, and a second list of all transitions from the DOWN state. For the list of transitions from the UP state, divide the list into sublists of size 30. We chose size 30 because it is the smallest sample size that is considered statistically significant. Allow the final sublist to be larger than 30 if need be. For the list of transitions from the DOWN state, do the same.

Step 4 - Perform Statistical Tests

The transition probability in a markov chain from state S1 to state S2 is a bernoulli random variable. Therefore, in the set of all transitions from S1, the number of occurrences from S1 to S2 is represented by a binomial distribution.

Perform this step on all six datasets.

In this experiment, the transition probabilities form UP and DOWN are unknown. We do, however know that the data is represented by a binomial distribution. For each sublist of transitions from UP, perform a binomial test with the set of probabilities {0.00, 0.01, 0.02, …, 1.00}, and record the results. Use the probability of an UP to UP transition as the probability being tested. When finished, do the same thing for each sublist of transitions from DOWN. Use the probability of a DOWN to DOWN transition as the probability being tested.

Step 5 - Analyze the Data

Perform this step on all six datasets.

The output of the binomial test is a probability of the likelihood of the data. Therefore, the higher the probability of the binomial test, the better the transition probability being tested fits the data.

If there is a transition probability P that fits the data for all sublists, then this shows that a markov chain with transition probability P would explain the data for a given transition in the independent, non-overlapping segments of bitcoin's price history that were tested.

If such a probability can be found for both the UP to UP transition and the DOWN to DOWN transition, then there is strong evidence that the price fluctuations of bitcoin can be explained meaningfully with a markov chain. The transition probabilities for UP to DOWN and DOWN to UP complement the previously mentioned transitions, and need not be explicitly calculated with a binomial test.

Experiment

Task #1

I cleaned the data using Microsoft Excel, and renamed the CSV file: "coindesk-bpi-USD-close_data-2013-01-01_2016-08-09.csv".

Task #2

I wrote a python script named "assignClasses.py". This script created the six datasets for the various d values, and assigned classes "UP" and "DOWN" accordingly. The six datasets were saved in a folder named "datasets".

Task #3

I wrote a python script called "recordTransitions.py". This script read in each CSV file from the "datasets" folder, identified and labeled all transitions, separated the transitions by their state of origin (ie. as transitioning from UP or transitioning from DOWN), divided the transitions into groups of 30, and recorded the transitions in a CSV file. The CSV files to which the transitions were recorded was stored in a folder named "transitions".

Task #4

I wrote a python script named “performBinomialTests.py”. This script read in each CSV from the “transitions” folder and for each division of 30 transitions performs a binomial test for probabilities {0.00, 0.01, 0.02, …, 1.00}. The results are recorded in a CSV and saved to a folder named “transitionProbabilities”.

Task #5

I wrote a python script named “analyzePValues.py”. This script creates a directory named “pValueResults” that has the same inner directory structure as “transitionProbabilities”. “analyzePValues.py” reads in each CSV from the “transitionProbabilities” folder, and for each CSV determined the range of probabilities for which the null hypothesis (ie. that the given probability explained the data) was not rejected. These results are recorded in a CSV in “pValueResults”.

Data Analysis

One of the central goals of this experiment is to determine a statistical measure to prove the validity of the results. To summarize the experiment simply, we are:

1. Taking the data and labeling each time period as "UP" or "DOWN" depending on if the price of bitcoin went up or down.
2. Calculating the probability of "UP" given some number of past probabilities (eg. probability of "UP" given that the previous day was a "DOWN", etc.) for the ENTIRE dataset.
3. Statistically test if the calculated probability of "UP" is internally consistent with the subsets within the dataset.

This creates a problem. The probability of "UP" calculated in step 2 above might be 0.5. However, when checking if this probability is internally consistent within the subsets of the dataset, one might find that the probability of "UP" for the first half of the time series data set is 0.8 and the probability of "UP" for the second half of the time series data set is 0.2. This would obviously warrant suspicion that the initial calculation of p(UP) = 0.5 is unusable as it is obvious that p(UP) is changing over the course of the time series. What, however, should one think if p(UP) for the first half was 0.45 and p(UP) for the second half was 0.55. This situation is less obvious. The problem here is that the calculated p(UP) for the subsets will (in all likelihood) indicate that the underlying process is non-stationary.

Conclusion

After consulting a friend who is getting a Ph.D. in Statistics, I have come to the conclusion that testing whether the deviations of p(UP) for the subsets are significantly different from p(UP) for the main set will be very difficult. I certainly do not know how to, and neither did my friend. There are a variety of tests that we could apply here, but the nature of this experiment would undermine their validity, as they rely on assumptions such as each data point being independent, etc. I originally began my bitcoin research with Markov chains because I thought that it would be simple. Now, I must end this line of investigation for the following reasons:

1. I will be unable to statistically validate my results for the aforementioned reasons.
2. There are other approaches to analyzing bitcoin that need to be tried.
3. The probability that p(UP) for the entire data set would NOT obviously differ from p(UP) of the subsets seems very slim to me. And if this difference is to be expected, then there is definitely no reason to continue. Even if I had the statistical know-how to do so, this phenomena would surely lead to the conclusion that Markov chains are no good for price prediction due to their non-stationary nature.

<u>Abstract</u>

No abstract necessary in the case, just read the conclusion.