

Anti-Streaky behavior in Currency Market

Teeranan Pokaparakarn (Ben)

Advisor: David Aldous

Introduction

Financial market has been extensively analyzed in the past decade, many try to explore and come with a model to take advantage of any non-random phenomenon in the market. This project does not attempt to test any statistical model but rather find some evidence of an anti-streaky behavior in financial market. The definition of anti-streaky in this case refer to roughly the idea that the market goes up and down too often even when it is trending. In this project, the price changes in the currency market are modeled simply as Bernoulli Trials. These sequences of ups and downs are then analyzed. The expectation and variance of the number of pairs of consecutive ups, pairs of consecutives down, triples of consecutives ups, as well as the patterns of up followed by down and then followed by up are analyzed. This kind of analysis which is the basis of this project is derived from the section 'Runs in the Stock Market' (Page 71) of the book Probability Tales by Grinstead, Peterson, and Snell.

Dataset

The data used in this project is the one-minute interval closing prices of the Euro-USD currency pair from January 1st 2008 to December 1st 2013. These 6 years of data is huge since we tracking the movement every minute. The reason for using the currency market is that it is often observed that the trends in currency market are usually stronger and last longer and intuitively one-minute price movements seem to be random. Hence, anti-streaky behavior may not be expected, and hence worth examining

The data set is then separated daily, each day the sequence of ups and downs is generated and the number of different patterns such as 'up-up', 'down-down', 'up-down-up', and 'up-up-up' are counted. To prevent noise from being a problem, a movement in the data is only counted as up or down if the price move by more than 3 pips (0.0003).

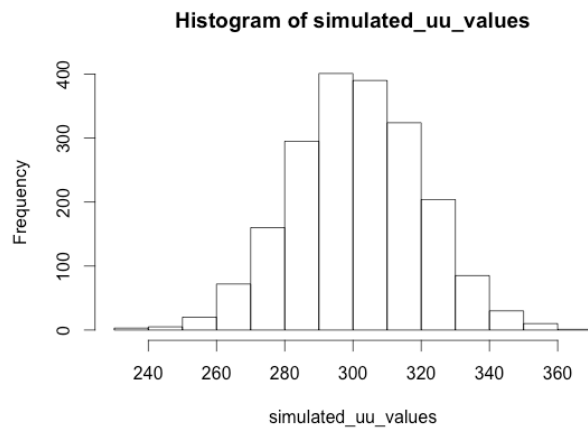


The figure above shows the boxplot of the number of daily price movements in our data in the year 2008. This give us a rough idea that on most days, we can get around 250 to 400 price movements which means 250 to 400 Bernoulli trials per day.

Methodology: Runs in Bernoulli Trials

In this project we are interested in the number of runs such as ‘up-up’, or ‘up-up-up’. Assuming that each price movement is a Bernoulli trial, the distribution of these runs can be found.

Let X_i be a random variable which equals to 1 if the i^{th} and the $(i+1)^{th}$ price movement are both up. Since $\{X_i\}$ is m-independent, a modified version of Central Limit Theorem applies (Probability Tales Page 73). To demonstrate this, a simulation of Bernoulli Trials is run and the result shown below also suggests that the number of pairs of consecutive ups is normally distributed.



This distribution also applies to other patterns such as ‘down-down’, ‘up-up-up’, or ‘up-down-up’.

Since the distribution is normal, we can count the number of these patterns in our data daily and calculated the z-values for each day. If the price were streaky, positive z-values for patterns such as ‘up-up’, ‘down-down’, or ‘up-up-up’ would be expected on most days. In other words, these patterns would appear more often than expected on most days. On the other hand, if the price were anti-streaky, more negative z-values would be an evidence. Another pattern that might be informative is ‘up-down-up’.

Mean and Variance Calculation

In order to calculate z-value, the mean and the variance of the counts of these patterns needs to be computed. For the mean, Markov chain method can be used. For instance, for ‘Up-Up’ pairs, a Markov chain which has four states {Up-Up, Up-Down, Down-Down, Down-Up} can be used to compute the stationary distribution for each state. Hence, an asymptotic mean of the counts can be obtained in this way. The same method applies for all the patterns.

For variance calculation, this can be done in the following way; suppose the goal is to find the variance of the counts of the pattern ‘Up-Up-Up’. Let A_i denotes the event ‘trials i , $i+1$, and $i+2$ are all Ups’. Let X denotes the count of how many of these events happen.

$$var X = \sum_{i=1}^m \sum_{j=1}^m [P(A_i \text{ and } A_j) - P(A_i)P(A_j)]$$

After careful counting and computing the probability of event as defined above, variance of the count of ‘Up-Up-Up’
 $= (n - 2)(p^3 - p^6) + 2(n - 3)(p^4 - p^6) + 2(n - 4)(p^5 - p^6)$;
 where n is the number of Bernoulli trials and p is the probability of ‘Up’.

After some calculation, below is the mean and variance of the patterns that are analyzed in this project. Let n and p defined as above, and $q = 1 - p$

'Up-Up': Mean = $(n-1)p^2$
Variance = $(n-1)(p^2 - p^4) + 2(n-2)(p^3 - p^4)$

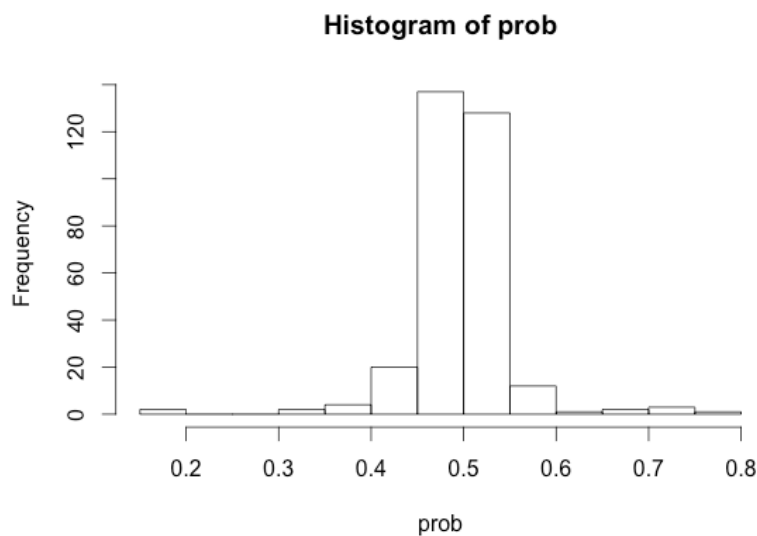
'Up-Up-Up': Mean = $(n-2)p^3$
Variance = $(n-2)(p^3 - p^6) + 2(n-3)(p^4 - p^6) + 2(n-4)(p^5 - p^6)$

'Down-Down': Mean = $(n-1)q^2$
Variance = $(n-1)(q^2 - q^4) + 2(n-2)(q^3 - q^4)$

'Up-Down-Up': Mean = $(n-2)p^2q$
Variance = $(n-2)(p^2q - p^4q^4) + 2(n-4)(p^3q^2 - p^4q^4)$

Parameter Estimation

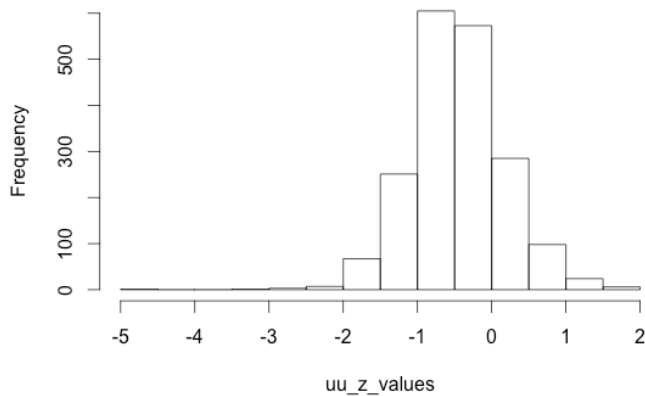
The only parameter that needs to be approximated in a Bernoulli trial is the probability of the price going up. In this case, since we are separating the data daily. The parameter, p, is estimated by the count of 'Up' movements divided the total count of price movement in a single day. The parameter p is estimated daily. To get an idea of what are these estimated values of p, below is a histogram of the estimated p in the year 2009.



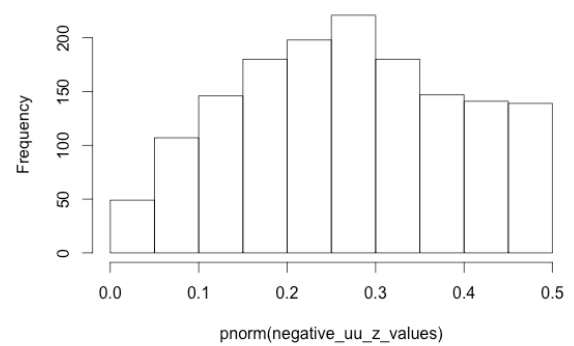
Result

The z-values from the year 2008 to 2013 of various patterns are combined and graphed together below. The histogram of z-values of streaky patterns (Up-Up), (Down, Down), and (Up, Up, Up) are shown on the left. On the right, since we are interested in the negative z-values, the histogram of p-value corresponding to the negative z-values in the left graphs is shown.

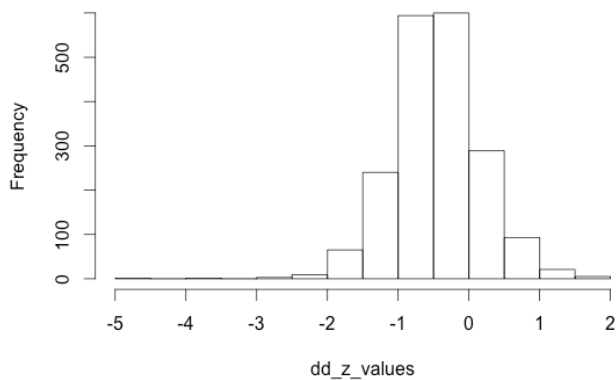
Histogram of uu_z_values



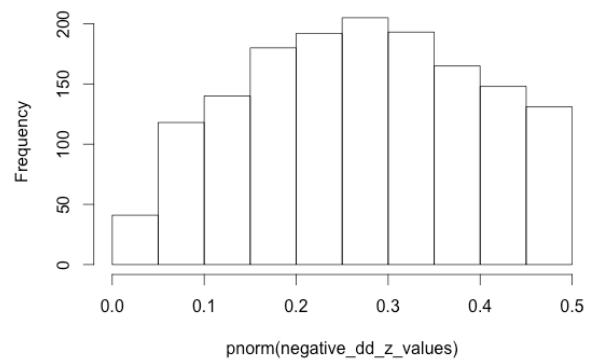
Histogram of pnorm(negative_uu_z_values)

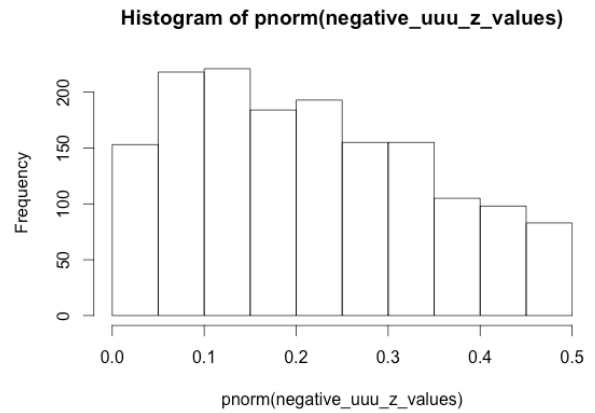
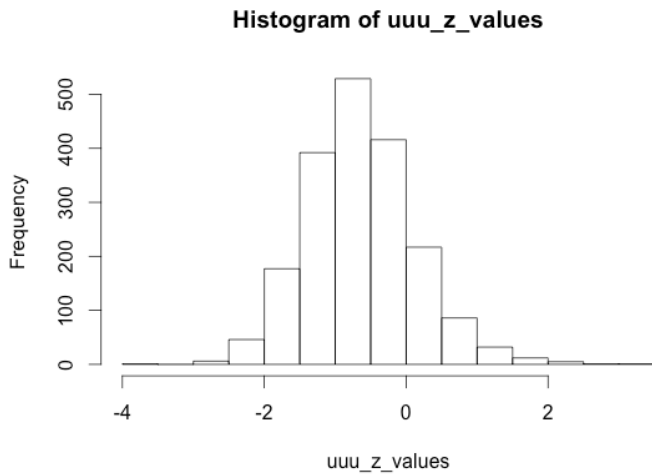


Histogram of dd_z_values



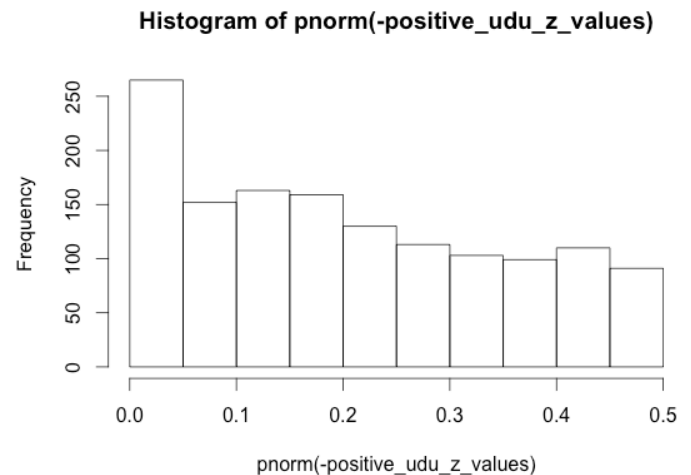
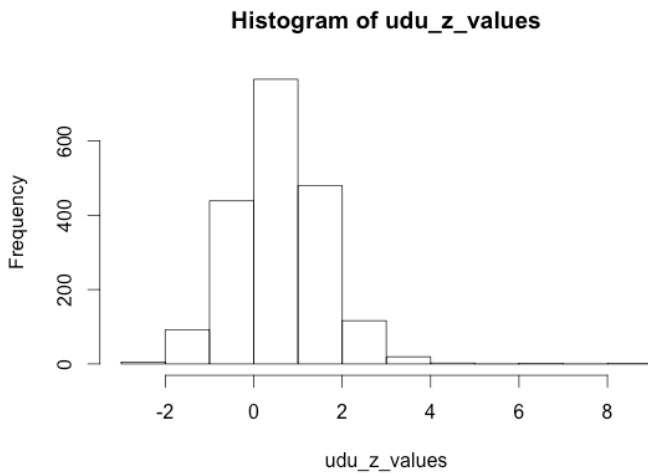
Histogram of pnorm(negative_dd_z_values)





Note: the z-values look normally distributed but centered on a negative value. For (Up,Up), 1511 z-values out of 1925 are negative. For (Up,Up,Up), 1567 z-values out of 1925 are negative.

Next, we examine an example of anti-streaky pattern (Up, Down, Up). In this case, the p-value plotted is for the positive z_values.



As expected, note that for anti-streaky pattern such as (Up, Down, Up), most z-values are positive (1388 out of 1925). Moreover, for many of these positive z-values, the p-value is quite small.

Discussion and Conclusion

As seen above, the z-values of pattern such 'Up-Up' looked normal but centered at a negative value. The fact that the z-values is consistently skewed or shifted toward the negative value as seen the graphs above implies that there are less 'Up-Up' pairs than would be expected assuming Bernoulli trials. The same observation can be made about 'Down-Down' pairs and the triples 'Up-Up-Up'. These 'streaky' patterns occur less often than expected. On the other hand, the pattern 'Up-Down-Up' seem to be centered around a positive z-value. Therefore, there is some evidence suggesting that the prices in this case might indeed be anti-streaky.

Although on most days, looking at the p-values, the z-values are not negative enough in the case of 'streaky' patterns such as 'Up-Up', 'Down-Down', and 'Up-Up-Up' to reject the hypothesis that this is just a due to chance. However, these z-values are still consistently negative for most of the days and in some days, the p-values are actually small enough (< 0.05) to reject the hypothesis. The anti-streaky behavior is further supported by the fact that patterns such as 'Up-Down-Up' mostly have positive z-values. It is hard to draw a conclusive claim that the market which in this case the Euro-USD currency pair is anti-streaky but this project shows that there is indeed evidence of anti-streaky behavior in the price movement assuming that price changes are Bernoulli trials.

Reference

1. "Probability Tales", Charles M. Grinstead, William P. Peterson, J. Laurie Snell (2011)

Data Source

1. <http://www.strategyquant.com/tickdatadownloader/>
(using Dukascopy Bank currency price data)