How humans take decisions: financial markets as a case of study

Author: Carlota Segura García
Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain.

Advisor: Josep Perelló

Abstract: Financial markets are a clear example of a choice-dealing situation in which people have to make decisions facing uncertainty and risk. This study aims to learn how people make decisions, in order to predict them, by analysing the data collected in a social experiment and using tools of information theory and statistics. It focuses on how people's own experience influence their next actions and what strategies are developed finding that both the market and the previous results influence decisions with a mutual information value of 0.045 ± 0.010 bits and 0.050 ± 0.010 bits respectively and that these two stochastic processes add non-redundant information to each other. Besides, in the experiment, people's memory holds for only one round. Finally, the 'toy' model tested gives a $55.200 \pm 0.016\%$ success ratio for the market's influence and a $53.496 \pm 0.016\%$ for the results influence analysis.

I. INTRODUCTION

Taking decisions is a very relevant issue in human's daily life: almost constantly people choose between different options, generally looking for the most beneficial result. But when humans have to take decisions there are many factors that can, conscious or unconsciously, affect their choice. For instance, their ability to understand the information, their previous experience or framing aspects among many other factors. This may lead them to develop basic strategies to address the problems they face. In the context of financial markets, economic theories have always tried to understand how traders make decisions thus bridging the gap between orders being placed in the market and the price dynamics. In the framework of classic economy, *utility theory* considers that humans are rational when making decisions and always bet for what they think will be most beneficial for them. Nevertheless, in the last decades alternative theories that introduce the traders psychology and also their ability to understand facts and information [3] have been proposed.

The experiment analysed took place the 14th and 15th of December 2013 at the Board Games Festival DAU Barcelona at Fabra i Coats-Fàbrica de Creació de Barcelona and was carried out by the Research Group Open Systems of the Department of Fundamental Physics of the University of Barcelona. It consisted of a game in which participants had to predict the market's daily action (rise or fall) for 25 consecutive rounds with maximum 30 seconds to answer per round. The market followed a daily-time evolution extracted from actual stock market series and depending on whether the participants' decision was right or not, they won or lost. Besides, participants had the opportunity at each round to consult some information available: the market's evolution averaged over the last 5 days and 30 days, the market's value evolution minute to minute of the last day, an expert's opinion (who was right 60% of times), the market's daily move (up or down) of the last 30 days and the tendency of some stock market's of the world, a rather typical basic information that market traders take into consideration. There were also four different modalities of game restricting different aspects such as the time available to answer or the information available. In the experiment participated 283 people (184 males and 99 females) of all ages from below 12 to 65 years old. The global success ratio was 53.43% and it has to be taken into account that, due to the actual market series used in the game, the market had a general tendency to rise of 55.69% (it is possible to play the game through the website, `mr-banks.net`).

For what it concerns to the data analysis point of view, each round can be characterized by the market's action, the player's decision and its result (right or wrong depending on whether the decision is equal to the market's movement or not). Thus, the data collected can be interpreted as three discrete temporal sequences or three stochastic binary processes: the market's evolution, the decision's path and the result series.

This study is developped in the framework of *information theory* in order to understand how the decision's path is affected by the other two sequences and so what amount of information is encoded in these processes and if it allows us to predict future decisions. In this sense, it should be noted that the underlying motivation in predicting decisions is not to know what the market's evolution is going to be but what the reactions to its movement or to the investing succes can be. Thus, in the following sections we study in detail the mutual information dependence between these pairs of processes alone and also their interdependence. Then we proceed to stablish the memory of a particular decision, i.e., how information is lost going backwards in time. Finally, with the results found a 'toy' model is tested with the aim to see if with this information it is actually possible to predict decisions.

II. QUANTIFYING THE INFLUENCES ON THE DECISION'S PATH: MUTUAL INFORMATION

Probability can be interpreted as the uncertainty about the occurrence or not of a certain event [1], that is to say, if the probability of a certain event is 1 (or 0) there is no such uncertainty: the event will always (never) take place. In this sense, a measure of the amount of uncertainty of a random variable X can be defined in terms of the *entropy of Shannon* $H(X)$

$$H(X) = -\sum_i p(x_i) \log_2(p(x_i)), \tag{1}$$

where the sum is extended over all the possible values x_i of the random variable X. As a convention, the logarithm is expressed with base 2 so all values are measured in *bits*.

Let us introduce too the *conditional entropy* $H(Y|X)$, the uncertainty on X provided Y occurs,

$$H(Y|X) = -\sum_y p(y|x) \log_2(p(y|x)). \tag{2}$$

If X and Y are two stochastic processes, a way to measure its mutual dependence is to compute its *mutual information* (MI). Taking these two entropies into account, the *mutual information* MI is defined as [1]

$$I(X,Y) = H(Y) - H(Y|X) =$$

$$= \sum_{x\in X, y\in Y} p(x,y) \log_2\left(\frac{p(x,y)}{p(x)p(y)}\right), \tag{3}$$

where $p(x,y)$ is the joint probability and $p(x)$ and $p(y)$, the marginal probabilities of the two random variables X and Y. MI is defined positive and, when equal to 0, variables are completely independent and, when equal to 1, they are perfectly correlated (for a deeper discussion see Appendix A). Actually, MI is a measure of the amount of information one random variable contains about another [2], so it will thus allow to find out how much the market's action and the player's result at a certain step tells us about the next decision.

First of all let us compute the MI involving the market's (M) previous action, at round '*n-1*', and the '*nth*' decision (D). Taking ino account the huge amount of information available in the game, imitating what the market has done would be the most logical and also simplest strategy to follow. As can be seen in Fig. 1, we are computing the MI of these two processes with the mentioned displacement since we are interested in the influence of the market's previous action on the decision.

The MI obtained for these two processes is (for a more detailed discussion of the errors see the Appendix B)

$$I(M_{n-1}, D_n) = 0.045 \pm 0.010 \text{ bits}.$$

Apart from imitating the market's movements, another usual way to proceed is by trial-error as people usually learn to choose by adapting its decisions to their success. For this reason we study the influence of the previous result on the next decision. Taking into account that when analysing what the previous result (R) has been ('*right*' or '*wrong*') we are not distinguishing what the previous decision was ('*up*' or '*down*'), the decision's path of the player is now expressed in terms of '*repeating*' or '*changing*' the decision (D) (for more clarity see Fig. 2).

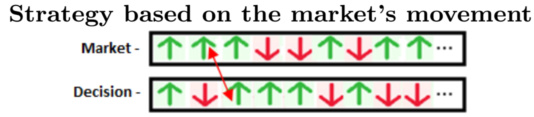

FIG. 1: A possible market's sequence and a possible decision's path. Green arrows pointing '*up*' mean that the market rises or that the player has made this prediction. Red arrows pointing '*down*' mean that it falls or it that the player has made this prediction. The inclined red arrow shows the displacement relation studied.

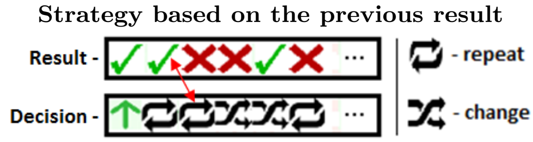

FIG. 2: A possible result sequence and a possible strategy followed: '*change*' or '*repeat*' the previous decision. The inclined red arrow shows the displacement relation studied. It has to be noted that the first decision has been discarded as it has no market previous action to be compared with.

The MI value obtained in this case is

$$I(R_{n-1}, D_n) = 0.050 \pm 0.010 \text{ bits}.$$

It can be seen that in both cases the value of MI is not equal to zero, which means that, as expected, the decision process is not independent of the market and the success process. To validate the importance of these results, we have computed the MI of the market with itself, $I(M_{n-1}, M_n) = 0.003 \pm 0.010$ bits, and also the MI of the player's own action, $I(D_{n-1}, D_n) = 0.005 \pm 0.010$ bits to see what information encode these processes alone and to better interpret the results obtained up to now. In these two last cases MI can be considered zero since its error covers this value. Taking a look to all the values found it can be said that neither the market nor the decision alone encode information about their own next action while the market and the result, separately, do have influence on the decision.

A. Interdependence between both pairs of sequences

We have studied the dependence of the decision's path on the market's action and the result sequence separately as if the market and the result affected independently the player. Nevertheless, it could be possible that these two processes were affected by each other and that when considering both influences more information was encoded.

In order to see if taking both processes into account increases the information, we compute the *conditional* MI (CMI)

$$I(X,Y|Z) = \sum_{x,y,z} p(x,y,z) \log_2 \frac{p(x,y,z)p(z)}{p(x,z)p(y,z)}. \quad (4)$$

Thus, we study the influence of the result on the market-decision comparison $I(M_{n-1}, D-n|R_{n-1})$ and also the influence of the market's action on the result-decision comparison $I(R_{n-1}, D_n|M_{n-1})$. As before, in these two values both the market and the result belong to round '*n-1*' while the decision belongs to the *nth* round. The CMIs obtained are

$$I(M_{n-1}, D_n|R_{n-1}) = 0.05 \pm 0.04 \text{ bits},$$

and

$$I(R_{n-1}, D_n|M_{n-1}) = 0.07 \pm 0.04 \text{ bits}.$$

Both values are positive which means that the market and the result sequence add non-redundant information. However, we shall also point that the market's behaviour adds more information.

B. Conditional Probabilities

To better understand the $I(M_{n-1}, D_n)$ and $I(R_{n-1}, D_n)$ values, the conditional probabilities for the decision being '*up*' or '*down*' knowing the market's previous action or the player's previous result and the marginal probabilities for the two versions of the decisions' path have been computed. It can be seen that these probabilities are statistically relevant in all cases compared to the marginal probabilities, as shown in Tab. I. This can make us think that these MI values obtained are too low, but the fact is that MI is a sum of terms weighted by its respective joint probabilities and so it can be interpreted as an average in which no special cases stand out.

Let us highlight an interesting result related to these probabilities. In the game proposed in the experiment, player's won or lost exactly the same amount (a 5% of their virtual 'money', starting from 1000) in the case they suceeded or failed respectively. Thereby, no matter what the market's actions were, if the predicition was right, the player won. In this sense, under the point of view of the *utility theory*, which considers decision-makers as rational [3], no difference between conditional probabilities given a particular action (rise, for instance) should be expected. However it is remarkable that $p(D_n^u|M_{n-1}^u)$ is much higher than $p(D_n^d|M_{n-1}^d)$. Thus, it can be said that there is a general tendency to follow the market when it is rising that has made player's think the market will keep on this upward trend for at least one more round.

Conditional probabilities	**Marginal probabilities**
Decision and market's events	
$p(D_n^u\|M_{n-1}^u) = 0.714 \pm 0.005$	
$p(D_n^d\|M_{n-1}^u) = 0.286 \pm 0.005$	$p(D_n^u) = 0.607 \pm 0.004$
$p(D_n^u\|M_{n-1}^d) = 0.469 \pm 0.006$	$p(D_n^d) = 0.393 \pm 0.004$
$p(D_n^d\|M_{n-1}^d) = 0.531 \pm 0.006$	
Decision and result's events	
$p(D_n^{rp}\|R_{n-1}^r) = 0.682 \pm 0.005$	
$p(D_n^{ch}\|R_{n-1}^r) = 0.318 \pm 0.005$	$p(D_n^{rp}) = 0.561 \pm 0.004$
$p(D_n^{rp}\|R_{n-1}^w) = 0.421 \pm 0.006$	$p(D_n^{ch}) = 0.439 \pm 0.004$
$p(D_n^{ch}\|R_{n-1}^w) = 0.579 \pm 0.006$	

TABLE I: Comparison between conditional probabilities and marginal probabilities. We show the conditional probabilities of the decision (D) being up, '*u*', or down, '*d*', in the *nth* round having the market (M) raised ('*u*') or fallen ('*d*') in the '*n-1*' round, the conditional probabilities of repeating ('*rp*') or changing ('*ch*') the decision (D) having the previous result (R) been right, '*r*', or wrong, '*w*', and the respective marginal probabilities for the decision being specifically up or down and for changing or repeating the previous decision.

III. LOSS OF INFORMATION

Up to now we have observed that the knowledge of the events of the previous round reduces uncertainty about the player's next decision, specially when considering particular cases. Such encodement of the information suggests that players have memory of the last round and that this influences their strategy. But now it can be asked whether this memory is wider or not, i.e., if, for instance, events that have taken place at round '*n-2*' can tell something about the *nth* decision.

In order to study how previous rounds are actually relevant on the decision, the MI of the market's action and the player's result at round '*n-2*' being the player's decision at the *nth* round have been computed obtaining

$$I(M_{n-2}, D_n) = 0.001 \pm 0.009 \text{ bits}$$

and

$$I(R_{n-2}, D_n) = 0.002 \pm 0.010 \text{ bits}.$$

We can observe that these values are not differentiable from 0 and, thus, in general no information about the next decision is encoded two rounds before.

Therefore, it can be said that only the previous round influences the player's decision, so its memory holds for only one round. Nevertheless, as said before, the MI is a thermalized value and in some special cases memory could be wider. For that reason, as done in Sect. II, the conditional probabilities considering the two previous rounds, $p(D_n|M_{n-1}, M_{n-2})$ and $p(D_n|R_{n-1}, D_{n-1}, R_{n-2})$, have been computed and are shown in Tab. II. It has to be noted that in the case of the result-decision comparison, the decision of the '*n-2*' round has been taken as a condition too since in this

Decision conditioned to		
Two previous market's action	**Two previous results and previous decision**	
$p(D_n^u\|M_{n-1}^u, M_{n-2}^u) = 0.709 \pm 0.006$	$\mathbf{p(D_n^{rp}\|R_{n-1}^r, D_{n-1}^{rp}, R_{n-2}^r) = 0.715 \pm 0.008}$	$p(D_n^{rp}\|R_{n-1}^r, D_{n-1}^{rp}, R_{n-2}^w) = 0.670 \pm 0.012$
$p(D_n^d\|M_{n-1}^u, M_{n-2}^u) = 0.291 \pm 0.006$	$\mathbf{p(D_n^{ch}\|R_{n-1}^r, D_{n-1}^{rp}, R_{n-2}^r) = 0.285 \pm 0.008}$	$p(D_n^{ch}\|R_{n-1}^r, D_{n-1}^{rp}, R_{n-2}^w) = 0.330 \pm 0.012$
$\mathbf{p(D_n^u\|M_{n-1}^d, M_{n-2}^u) = 0.493 \pm 0.008}$	$p(D_n^{rp}\|R_{n-1}^w, D_{n-1}^{rp}, R_{n-2}^r) = 0.412 \pm 0.010$	$\mathbf{p(D_n^{rp}\|R_{n-1}^w, D_{n-1}^{rp}, R_{n-2}^w) = 0.550 \pm 0.013}$
$\mathbf{p(D_n^d\|M_{n-1}^d, M_{n-2}^u) = 0.507 \pm 0.008}$	$p(D_n^{ch}\|R_{n-1}^w, D_{n-1}^{rp}, R_{n-2}^r) = 0.588 \pm 0.010$	$\mathbf{p(D_n^{ch}\|R_{n-1}^w, D_{n-1}^{rp}, R_{n-2}^w) = 0.450 \pm 0.013}$
$p(D_n^u\|M_{n-1}^u, M_{n-2}^d) = 0.716 \pm 0.007$	$\mathbf{p(D_n^{rp}\|R_{n-1}^r, D_{n-1}^{ch}, R_{n-2}^r) = 0.526 \pm 0.014}$	$\mathbf{p(D_n^{rp}\|R_{n-1}^r, D_{n-1}^{ch}, R_{n-2}^w) = 0.715 \pm 0.010}$
$p(D_n^d\|M_{n-1}^u, M_{n-2}^d) = 0.284 \pm 0.007$	$\mathbf{p(D_n^{ch}\|R_{n-1}^r, D_{n-1}^{ch}, R_{n-2}^r) = 0.474 \pm 0.014}$	$\mathbf{p(D_n^{ch}\|R_{n-1}^r, D_{n-1}^{ch}, R_{n-2}^w) = 0.285 \pm 0.010}$
$\mathbf{p(D_n^u\|M_{n-1}^d, M_{n-2}^d) = 0.447 \pm 0.008}$	$p(D_n^{rp}\|R_{n-1}^w, D_{n-1}^{ch}, R_{n-2}^r) = 0.408 \pm 0.014$	$\mathbf{p(D_n^{rp}\|R_{n-1}^w, D_{n-1}^{ch}, R_{n-2}^w) = 0.350 \pm 0.011}$
$\mathbf{p(D_n^d\|M_{n-1}^d, M_{n-2}^d) = 0.553 \pm 0.008}$	$p(D_n^{ch}\|R_{n-1}^w, D_{n-1}^{ch}, R_{n-2}^r) = 0.592 \pm 0.014$	$\mathbf{p(D_n^{ch}\|R_{n-1}^w, D_{n-1}^{ch}, R_{n-2}^w) = 0.650 \pm 0.011}$

TABLE II: Conditional probabilites of the decision considering the last two rounds. Shows the conditional probabilities of the decision (D) depending on the market's (M) or the result's (R) two previous events. The superindices '*rp*' mean '*repeat*' decision, '*ch*' means '*change*' decision, '*r*' means that the prediction was '*right*' and '*w*' that it was '*wrong*'. The values in bold are those statistically relevant compared to values shown in Tab. I.

case we are interested in knowing how the player adapts at each step its strategy and thus if, for example, it tends to repeat its decision when it has been successful.

We can observe that, when comparing the probabilities that represent the same event in the *nth* and '*n-1*' rounds, in some cases (values in bold) these conditional probabilities are statistically relevant from the ones shown in Tab. I while in some other cases they are not. That is to say that, for example, we have compared $p(D_n^u|M_{n-1}^d, M_{n-2}^u)$ with $p(D_n^u|M_{n-1}^d)$ and $p(D_n^{rp}|R_{n-1}^r, D_{n-1}^{ch}, R_{n-2}^w)$ with $p(D_n^{rp}|R_{n-1}^r)$.

It is remarkable that the probability of repeating the decision when the result has been correct, increases when this situation takes place two consecutive times compared to when it only happens once. In a similar way, the probability of changing the decision when the player has failed to predict also increases when these events take place twice. Therefore, it seems like there is a tendency of the players to adapt their decisions depending on their success.

On the other hand, the probability of following the market's rising trend of the two previous rounds is not differentiable from that of considering only the previous round. Contrarily, the probability of following the market's tendency when it has fallen twice increases with respect to considering only the previous movement. This facts can be understood as there is some distrust to believe that the upward trend will keep for more rounds, although this probability is still high, while if the tendency is downwards the reliance on this trend increases.

IV. TESTING A 'TOY' MODEL

Troughout the study, we have observed that information about the player's decisions is indeed encoded in the market's different actions and in the player's own results. Besides, in the last section we have seen that in terms of MI, only what has happened in the previous round influences the next decision.

In order to see if taking all these facts into account we are capable of predicting what the decisions are, and so how people take decisions, we propose a 'toy' model using the conditional probabilities shown in Tab. I. However, as the data available is limited, in order to test the model in a more rigorous way, the conditional probabilities have been computed using only half of the population's data (selected randomnly). Then, these conditional probabilities have been used to predict what the decisions of the other half of the population are in the two different analysis developed throghout this study, the market-decision and the result-decision comparisons, according to the actual moves of the market and results of the players in each game. This process has been repeated 1000 times in each analysis in order to obtain the average ratio of success and its standard deviation.

In the case of the analysis of the market's influence the success ratio obtained has been $55.200 \pm 0.016\%$. Similarly, for the result's influence the success ratio obtained has been $53.496 \pm 0.016\%$. Both are a little bit higher than 50% (throwing a coin). It has to be highlighted, though, that we have considered that all player's make decisions in the same way, i.e., we are not distinguishing between those who follow strategies from those who play more randomnly.

This 'toy' model can still be improved since we could add the fact that both the market and the previous result influence at the same time the player's decision. Besides, although it has been seen that the memory stands only for one round in terms of the MI, some of the conditional probabilities going to two rounds before are still relevant and, thus, the model could be more detailed and implement these probabilities too.

V. CONCLUSIONS

Information theory has applications in many different fields such as neurobiology, thermal physics or quantum computing [2]. In this study it has proved to be a useful alternative to merely statistics as it has allowed us to extract more information about what influences decisions and also to complement the conditional probabilities obtained.

We have observed that both the market's action and the results obtained influence the player's decisions since both processes encode information about the latter as their MI values, $I(M_{n-1}, D_n)$ and $I(R_{n-1}, D_n)$, are different from zero. Besides, these two processes, the market's and the reult's one, add non-redundant information to each other, which means that, as could be expected, player's decisions are influenced at the same time by the market's movements and by their success.

In general, the player's memory stands only for one round in both analysis since the value of MI drops abruptly when considering the two previous rounds compared to the value obtained considering only to the previous round and it cannot be distinguished from zero. Nevertheless, some of the conditional probabilities computed for the two previous rounds shown in Tab. II are still statistically relevant compared to those computed with only one condition shown in Tab. I.

When looking in detail these sets of conditional probabilities for the market-decision comparison looking for the basic strategy of imitating the market, it is remarkable the great difference between the tendency to follow the market when it has risen than when it has fallen. This fact shows that player's have a relative preference for the market to rise and thus they are not strictly thinking in terms of only benefit and this contradicts the assumptions of rationality of utility theory. On the other hand, when considering two rounds, we observe that the probabilities of following the market's clear upward and downward tendencies do not vary or increase considerably compared to going to one step but rather they are mantained. This reflects a certain resistance of player's to trust in long trends.

On the other hand, when looking the sets of conditional probabilities for the result-decision comparison it we observe that there is a clear strategy that consists in betting by trial-error, i.e., repeating the decision when being successful and changing it when it has failed, that is increased when considering the two previous rounds.

Testing the fact that memory holds for one round and implementing the conditional probabilities shown in Tab. I, we are able to predict player's decisions in more than 50% of the cases. Nevertheless, this 'toy' model could be improved using the statistically relevant conditional probabilities when gonig to two rounds before and also trying to distinguish different ways of deciding.

VI. APPENDIX

A. Mutual Information

In order to verify that the MI algorithm has been correctly implemented, two programs have been designed. The first one generates randomnly a market sequence and a decision path so that when computing MI the expected value is 0 as they are independent. The second program generates a market sequence randomnly and a decision path following exactly what the market has done in the previous round. In this case the expected value is 1 as they are perfectly related. Since in both cases these values have been obtained with a discrepancy of just the $0,01\%$, it can be considered that MI algorithm used throughout the study is coorect.

B. Errors

As mentioned in the Introduction, all three stochastic processes (market, result and decision's sequences) are binary as only two outcomes are possible: '*up*'/'*down*' or '*right*'/'*wrong*'. Therefore, the error for the empirical probabilities computed from the data has been

$$\delta p = z\sqrt{\frac{p(1-p)}{N}}, \tag{5}$$

where z is the percentile for a standard normal distribution (in this case $z = 1$ has been used, which implies a 68% confidence level), p is the empirical probability and N is the sample size (which can be different for different joint and conditional probabilities). Proceeding from this, all errors throughout the study have been computed by error propagation.

Acknowledgments

I want to thank my advisor, Josep Perelló, who has helped me a lot throughout the study and also to Mario Gutiérrez-Roig and Isabelle Bonheure who have been part of the team too for their ideas, advices and dedication in all the meetings we have had.

[1] Athanasios Papoulis and S. Unnikrishna Pillai, *Probability, Random Variables and Stochastic Processes*, 4th. ed. (Mc Graw Hill, New York, 1984).

[2] Thomas M. Cover and Joy A. Thomas, *Elements of information theory*, 2nd ed. (John Wiley and sons, New York, 1991).

[3] Tversky, Amos, and Daniel Kahneman. "The framing of decisions and the psychology of choice." Science 211.4481 (1981): 453-458.