Ç.Ü. Sosyal Bilimler Enstitüsü Dergisi, Cilt 22, Sayı1, 2013, Sayfa 501-512

Estimating Probability of Session Returns for Istanbul Stock Exchange 100 Index as Markov Chain Process

Süleyman Bilgin KILIÇ*

ABSTRACT
In this study, I modeled session returns for the Istanbul Stock Exchange 100 (ISE100) index as the eight discrete state Markov chain process in order to estimate session returns of the ISE100 index. The model provides valuable signals to the investors about short run selling and buying investment strategies.

Keywords: ISE 100, Stock Returns, Markov chains, Conditional probability

1. Introduction

Objective of this study is to examine and explain directional movements of returns for the ISE100 index as eight discrete states Markov chain process in order to provide information to the investors for their short term selling and buying investment strategies. Many important studies modeled and analyzed stock returns as Markov chains structure. Empirical results of these studies provide valuable information to the investors. In the study of Flietz and Bhargava (1973) empirical results for the vector process model suggest that price movements appear to be described by a first- or higher-order non-stationary Markov chain. Tests also indicate that the vector-process Markov chain is heterogeneous. Empirical results for the individual-process Markov chain model suggest that an individual stock has a short-term memory with respect to daily price relatives. Ryan (1973) showed that both the relative strength of a security in market and the nature of its successive price movements may be interpreted within the framework of Markov theory in such a way as to provide useful information to the portfolio manager. Turner, Startz and Nelson (1989) used a Markov model of heteroskedasticity, risk and learning in the stock market. The estimates indicate that agents are consistently surprised by high variance periods, so there is a negative correlation between movements in volatility and in excess returns. In the Markovian study of Mcquinn and Thorley (1991) annual real returns were shown to exhibit significant nonrandom walk behavior in the sense that low (high) returns tend to follow runs of high (low) returns in the postwar period. Hamilton and Susmel (1994) examined the U.S. weekly stock returns, allowing the parameters of an ARCH process to come from one of several different regimes, with transitions between regimes governed by an unobserved Markov chain. They estimated models with two to four regimes in which the latent innovations come from Gaussian and Student *t* distributions. In the study of Mills and Jordonov (2003) the returns showed evidence of nonlinearity and non-normality, so conventional autocorrelation analysis was supplemented by the use of a Markov chain technique. They found evidence that the returns of the two smallest and four largest size portfolios were predictable, but only the two largest were predictable in

*Doç.Dr., Çukurova Üniversitesi, İİBF, Ekonometri Bölümü, Yöneylem Anabilim Dalı, sbilgin@cu.edu.tr

the direction suggested by the bubbles and fads alternatives to the random walk hypothesis. Kanas (2003) suggested that Markov regime switching model is the most preferable non linear empirical extension of the present value model for out-of-sample stock return forecasting. Kılıç (2005) found it is not possible earn above average return in the long run. In the study of Idolor (2010) eight securities were selected from the banking sector of the Nigerian Stock Exchange by defining a set of three states (rise, drop, and stable) for the Markovian process. Results showed that Markov chains did not give a precise prediction of the direction in which prices were headed in the short run.

Some other important studies combined Markov chain model with the Monte Carlo simulation technique, GARCH time series, and/or Bayesian forecast models. Tsionas (2000) illustrated the computation of marginal likelihoods and Bayes factors when Markov Chain Monte Carlo had been used to produce draws from a model's posterior distribution. Models included a normal finite mixture, a GARCH and a Student *t* -model as alternative models for the Standard and Poor's stock returns. In the study of Griffin and Steel (2006) the algorithm was based on Markov chain Monte Carlo (MCMC) methods and they used a series representation of Lévy processes. Their application to stock price data showed the models performed very well, even in the face of data with rapid changes, especially if a superposition of processes with different risk premiums and a leverage effect was used. Greyserman, Jones and Strawderman (2006) contributed to portfolio selection methodology using a Bayesian forecast of the distribution of returns by stochastic approximation. They carried out a numerical optimization procedure to maximize expected utility using the MCMC samples from the posterior predictive distribution. This model resulted in an extra 1.5 percentage points per year in additional portfolio performance which is quite a significant empirical result. Chen and So (2006) proposed a threshold heteroscedastic model which integrates threshold nonlinearity and GARCH-type conditional variance for modeling mean and volatility asymmetries in financial markets. They showed higher average volatility and more persistent volatility when bad news arrives. Guidolin and Timmermann (2007) characterized equilibrium asset prices under adaptive, rational and Bayesian learning schemes in a model where dividends evolve on a binomial lattice. They investigated restrictions on prior beliefs under which Bayesian and rational learning lead to identical prices and show how the results can be generalized to more complex settings where dividends follow either multi-state i.i.d. distributions or multi-state Markov chains. Zhang and King (2008) presented a MCMC algorithm to estimate parameters and latent stochastic processes in the asymmetric stochastic volatility (SV) model, in which the Box-Cox transformation of the squared volatility follows an autoregressive Gaussian distribution and the marginal density of asset returns has heavy-tails. They found that when their model and its competing models were applied to daily returns of another five stock indices, in terms of SV models, the Box-Cox transformation of squared volatility is strongly favored against the log-transformation for the five data sets. Lin, Wang, and Lin, Wang and Tsai (2009) proposed a hidden Markov switching moving average model (MS-MA model) to extend the moving average model when the dynamic process of stock returns is predictable. They showed that the dynamic process of stock returns exhibits MS-MA property, meaning the moving averages of stock returns are correlated. Liu (2011) examined a continuous- time intertemporal consumption and portfolio choice problem under ambiguity, where expected returns of a risky asset follow a hidden Markov chain. He found that continues Bayesian revisions under incomplete information generate ambiguity-driven hedging demands that mitigate intertemporal hedging demands.

In this study first we modeled session returns for the Istanbul Stock Exchange 100 (ISE100) index as Markov chains process to examine and explain directional movements of returns. Modeling returns as discrete categorical states help us to see and explain upward and downward directional movements and also to measure the dimension of these movements more precisely. Then, we calculated the total number of occurrence of the transitions of states from the present session to the next session for the period considered. Third, we calculated conditional probabilities of next return states given the value of present state of returns and obtained one step transition probability matrix. By using this matrix we calculated probabilities of expected return for each state. We also examine that whether the transition probability matrix follow property of a regular ergodic Markov chain structure in the long run. We found that Markov chains process provides valuable signals to the investors.

The rest of this article is organized as follows. Section 2 includes the data selection, methodology and empirical results; section 3 concludes the article and discusses some future research perspectives.

2. The Sample, Methodology and Results

2.1. The Sample

The sample data covers 10357 session closing values of the Istanbul Stock Exchange 100 (ISE100) index between the period of January 04, 1988-April 04, 2012. The data were obtained from the electronic data delivery system of the Istanbul Stock Exchange (http://www.imkb.gov.tr).

The daily session returns (R_t) are computed as a percentage change of the ISE100 index session closing values (P_t); $R_t = (P_t - P_{t-1})/P_{t-1}$. Here, *t* represents the sessions (*t*=1…10356). The average (expected) return is calculated (μ_R) as approximately 0.108% with a standard deviation of (σ_R) 2.28% for the period considered. The standard deviation is extremely high in comparison to expected return (approximately 23 times of the expected return). So, the ISE100 index return exhibits high dispersion, implying an enormous risk for the investors.

2.2. The Methodology and Results

We can model the ISE100 returns as a Markov chains process that conditional probability of any next future return state (S_j^{t+1}) depends only on the present return of the state (S_i^t) and is independent any other states of the past returns; $P\{S_j^{t+1} | S_i^t\}$.

As stated previously, modeling the ISE100 index returns as the discrete categorical states Markov chain process help us to see and explain directional movements and also to evaluate the dimension of these movements more precisely. Thus, session returns are transformed into eight equal discrete categorical interval of states, from high loss (negative return) to positive high return according to function 1 below;

$$(S_i^t) = \begin{cases} i=1, \; if \;\; R_t < -0.0171 \\ i=2, \; if \;\; -0.0171 \le R_t < -0.0114 \\ i=3, \; if \;\; -0.0114 \le R_t < -0.0057 \\ i=4, \; if \;\; -0.0057 \le R_t < 0 \\ i=5, \; if \;\; 0 \le R_t < 0.0057 \\ i=6, \; if \;\; 0.0057 \le R_t < 0.0114 \\ i=7, \; if \;\; 0.0114 \le R_t < 0.0171 \\ i=8, \; if \;\; 0.0171 \ge R_t \end{cases} \tag{1}$$

The total number of transitions, occurring from the present session to the next session from states S_i to S_j , are calculated for the period considered in Table 1[1]:

Table 1: Number of occurrence of the transitions from states S_i to S_j

	Next sessions								
Present Sessions	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	Row total
S_1	320	111	142	139	125	119	89	231	**1276**
S_2	124	77	98	122	123	81	58	103	**786**
S_3	138	105	176	194	239	152	87	127	**1218**
S_4	154	124	220	346	282	240	144	159	**1669**
S_5	136	129	186	329	309	229	130	179	**1627**
S_6	114	96	153	232	238	223	122	164	**1342**
S_7	80	56	93	122	142	130	99	149	**871**
S_8	210	88	150	186	169	168	142	453	**1566**

We can easily compute the eight states (8x8), one step (one session) conditional transition probability matrix $P\left\{S_j^{t+1} \middle| S_i^t\right\}$ in Table 2, from state *i* to *j* by dividing the row elements by row total in Table 1.

[1]Categories of states and number of transitions of states formulated and calculated in Excel by using Excel IF function.

Table 2: One step conditional transition probability matrix

	Next sessions							
Present Sessions	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
S_1	0.2508	0.0870	0.1113	0.1089	0.0980	0.0933	0.0697	0.1810
S_2	0.1578	0.0980	0.1247	0.1552	0.1565	0.1031	0.0738	0.1310
S_3	0.1133	0.0862	0.1445	0.1593	0.1962	0.1248	0.0714	0.1043
S_4	0.0923	0.0743	0.1318	0.2073	0.1690	0.1438	0.0863	0.0953
S_5	0.0836	0.0793	0.1143	0.2022	0.1899	0.1407	0.0799	0.1100
S_6	0.0849	0.0715	0.1140	0.1729	0.1773	0.1662	0.0909	0.1222
S_7	0.0918	0.0643	0.1068	0.1401	0.1630	0.1493	0.1137	0.1711
S_8	0.1341	0.0562	0.0958	0.1188	0.1079	0.1073	0.0907	0.2893

In Table 2, the session returns are assumed to be a stochastic process with eight discrete state spaces $\{S_1,..,S_8\}$ with Markov chain structure. For example, when the ISE100 index return in state S_5 in the present session, conditional probability of it will be going to S_3 in the next session is $P(S_3 | S_5) =$ 11.43%. Similarly, conditional probability of passing from state S_1 to S_4 is $P(S_4 | S_1) =$ 10.89%.

The one step transition matrix (Table 2) shows the following property of a regular ergodic Markov chain structure that $\underset{t\to\infty}{Lim.}\ P_{ij}^t = \pi_j$, here π_j 's are the steady state probabilities and this limit is independent of *i*. The π_j 's satisfy the following steady state conditions;

$$\text{for } \pi_j > 0,\ \sum_{j=1}^{M} \pi_j = 1, \text{ and } \pi_j = \sum_{i=1}^{M} \pi_i P_{ij}. \qquad (2)$$

We can find the value of π_j 's (steady state probabilities) that is independent of the initial probability distribution after a sufficiently large number of transitions. As *t* becomes larger, the values of the P_{ij}^t moves to fixed limit and each probability vector tend to become equal for all values of *i*. Thus, each of the eight rows of P_{ij}^t has identical probabilities;

Table 3: Six steps probability matrixes

P_{ij}^{1}

0.251	0.087	0.111	0.109	0.098	0.093	0.070	0.181
0.158	0.098	0.125	0.155	0.156	0.103	0.074	0.131
0.113	0.086	0.144	0.159	0.196	0.125	0.071	0.104
0.092	0.074	0.132	0.207	0.169	0.144	0.086	0.095
0.084	0.079	0.114	0.202	0.190	0.141	0.080	0.110
0.085	0.072	0.114	0.173	0.177	0.166	0.091	0.122
0.092	0.064	0.107	0.140	0.163	0.149	0.114	0.171
0.134	0.056	0.096	0.119	0.108	0.107	0.091	0.289

P_{ij}^{2}

0.146	0.077	0.116	0.148	0.144	0.121	0.082	0.165
0.130	0.077	0.118	0.160	0.156	0.127	0.083	0.150
0.121	0.078	0.120	0.165	0.162	0.131	0.083	0.141
0.116	0.077	0.120	0.168	0.163	0.133	0.084	0.138
0.115	0.076	0.119	0.169	0.163	0.133	0.084	0.140
0.115	0.076	0.118	0.166	0.162	0.134	0.085	0.144
0.118	0.074	0.116	0.161	0.158	0.132	0.086	0.155
0.130	0.073	0.113	0.151	0.146	0.125	0.086	0.177

:

P_{ij}^{6}

0.123	0.076	0.118	0.161	0.157	0.130	0.084	0.151
0.123	0.076	0.118	0.161	0.157	0.130	0.084	0.151
0.123	0.076	0.118	0.161	0.157	0.130	0.084	0.151
0.123	0.076	0.118	0.161	0.157	0.130	0.084	0.151
0.123	0.076	0.118	0.161	0.157	0.130	0.084	0.151
0.123	0.076	0.118	0.161	0.157	0.130	0.084	0.151
0.123	0.076	0.118	0.161	0.157	0.130	0.084	0.151
0.123	0.076	0.118	0.161	0.157	0.130	0.084	0.151

Table 3 gives transition probabilities going from one to six sessions respectively. From the last six step steady state probabilities matrix in Table 3, (P_{ij}^{6}) we can see that there is a limiting probability that the return states will be in steady state condition after six steps (t=6 session or three work days), independent of the present initial state.

By using above probability matrixes we can calculate expected capital gain or loss $E(C_i^t)$ for investors in the ISE after *t* session given the present state *i* as:

$$E(C_i^t) = C_i^{t-1}\left(1 + (P_{ij}^t.\mu_i)\right) \tag{3}$$

In equation 3, μ_i represents the vector of the mean returns of states. Table 4 gives the means and standard deviations of each state.

Table 4: Mean and standard deviation of returns of states

States	Return range	Mean (μ_i)	Std.dev. (σ_i)
S_1	$R_t < -0.0171$	-0.0328	0.0173
S_2	$-0.0171 \le R_t < -0.0114$	-0.0141	0.0016
S_3	$-0.0114 \le R_t < -0.0057$	-0.0084	0.0016
S_4	$-0.0057 \le R_t < 0$	-0.0027	0.0017
S_5	$0 \le R_t < 0.0057$	0.0028	0.0016
S_6	$0.0057 \le R_t < 0.0114$	0.0083	0.0016
S_7	$0.0114 \le R_t < 0.0171$	0.0140	0.0016
S_8	$0.0171 \ge R_t$	0.0326	0.0169

The expected capital gain or loss $E(G_i^t)$ net of the costs for the future *t* sessions, given the present state *i* can be calculated as:

$$E(G_i^t) = \begin{cases} E(C_i^t) - E(C_i^t)(0.10) - (C_i^0)(0.00175) - E(C_i^t)(0.00175), & if \;\; E(C_i^t) \ge C_i^0 \\ E(C_i^t) - (C_i^0)(0.00175) - E(C_i^t)(0.00175), & if \;\; E(C_i^t) < C_i^0 \end{cases} \tag{4}$$

In equation 4, C_i^0 represents the initial amount of capital invested by an investor in state *i* of the ISE index. The constant term, 0.10, represents the capital gain tax rate in the ISE, and the term 0.00175 represents buying and selling costs respectively[2]. We can calculate expected return $E(R_i^t)$ after the *t* sessions, given the present state *i* by equation 5,

$$E(R_i^t) = \left[(E(G_i^t) - C_i^0)/C_i^0\right] \tag{5}$$

[2] We take selling and buying commission of Garanti Bank for investment range of 5000-25.000TL.

Table 5: Calculated expected return after the *t* session given the present state *i*

Present state	t=1	t=2	t=3	t=4	t=5	t=6	t=7	t=8	t=9
S_1	-0.0062	-0.0056	-0.0045	-0.0034	-0.0025	-0.0015	-0.0005	0.0005	0.0015*
S_2	-0.0049	-0.0042	-0.0032	-0.0022	-0.0012	-0.0002	0.0007	0.0017*	0.0027
S_3	-0.0041	-0.0033	-0.0024	-0.0014	-0.0004	0.0005	0.0015*	0.0025	0.0035
S_4	-0.0033	-0.0025	-0.0015	-0.0006	0.0004	0.0014*	0.0024	0.0034	0.0044
S_5	-0.0026	-0.0017	-0.0007	0.0003	0.0012*	0.0022	0.0032	0.0042	0.0052
S_6	-0.0018	-0.0007	0.0002	0.0012*	0.0022	0.0032	0.0042	0.0052	0.0062
S_7	-0.0002	0.0011*	0.0021	0.0031	0.0041	0.0051	0.0061	0.0071	0.0081
S_8	0.0015*	0.0031	0.0042	0.0052	0.0062	0.0072	0.0082	0.0092	0.0102

*Future expected net session returns above the overall average session return (0.1084%) of ISE 100 index.

Table 5 gives the calculated future expected net session return $E(R_i^t)$ of the ISE index after buying and selling costs and the capital gain tax, after the *t* sessions, given the present state *i*. Here, one can see that if the present state is S_1 (high loss state) it is expected that return will be above the overall average session return (0.108%) of ISE 100 index after nine steps. If the present state is S_7 it is expected that return will be equal or above overall average return after one step. This means that if an investor buys the ISE 100 index and the present state is S_1 he or she has to wait nine sessions without selling in order to realize net return which is above the overall average session return. Similarly, if the present state is S_7 he or she has to wait one session only.

It is expected that an investor will be realize return which is equal to average session return of ISE 100 index in the long run. Because, as stated before there is a limiting probability that the return states will be in steady state condition after six steps (t=6 session or three work days), independent of the present initial state. If we multiply steady state probability vector (π_j) with the vector of the mean returns of states (equation 6), we can get long run average return.

$$\mu_R = \mu_i.\pi_j = \begin{bmatrix} 0.123 & 0.123 & 0.076 & 0.118 & 0.161 & 0.157 & 0.130 & 0.084 \end{bmatrix} \begin{bmatrix} -0.0328 \\ -0.0141 \\ -0.0084 \\ -0.0027 \\ 0.0028 \\ 0.0083 \\ 0.0140 \\ 0.0326 \end{bmatrix} = 0.001084 \qquad (6)$$

These empirical results generally suggest that when the ISE index is in the high positive return states (S_6, S_7 and S_8) it is expected that investors will realize a return which is above overall average session return much more rapidly. When the ISE index is in the high loss states (S_1 and S_2) investors may also hold short selling strategy.

3. Conclusion

This study provides information about investment opportunities of the ISE 100 index for their short term (nine sessions or approximately five days) buying and selling strategies given the present state. If the present state is known (e.g. S_6, S_7 and S_8) it is possible to earn above overall average return in the very short run. Similar further analysis can be performed by considering returns of smaller time intervals such as returns of ten or five minutes. Hence, small time intervals may provide more investment opportunities.

Result of the study is also consistent with the weak form of the efficient market hypothesis[3] in the long run. Because, the ISE index returns go to the steady state condition after six steps an investor's expected return will be equal to the long run average return of the ISE index regardless of the present state. The weak form efficient market hypothesis asserts that it is not possible earn above average return in the long run. However, according to the results of this study, the efficient market hypothesis may not hold in the very short run. In other words, when the very short time interval is considered, information in the past data is not fully reflected in present prices.

Similar further analysis can also be performed for returns of individual common stocks and other investment instruments such as gold and foreign exchange returns.

---

[3] Efficient Market Hypothesis (EMH) Much of the theory on these subjects can be traced to French mathematician Louis Bachelier whose Ph.D. dissertation titled "The Theory of Speculation" (1900). EMH evolved by Eugene, F.F. (1965) who proposed three forms of the efficient market hypothesis: (1) The "Weak" form asserts that all past market prices and data are fully reflected in securities prices. In other words, it is not possible earn above average return in the long run. (2) The "Semi-strong" form asserts that all publicly available information is fully reflected in securities prices. In other words, fundamental analysis is of no use. (3) The "Strong" form asserts that all information is fully reflected in securities prices. In other words, even insider information is of no use.

References

Chen, C. W. S. & So, M. K. P. (2006). On a threshold heteroscedastic model *International Journal of Forecasting*, Volume 22, Issue 1, January–March 2006, pp 73-89.

Eugene F.F. (1965). Random Walks in Stock Market Prices, *Financial Analysts Journal*, September/October.

Flietz, B. D. & Bhargava, T. N. (1973). The Behavior of Stock-Price Relatives-A Markovian Analysis, *Operations Research, Vol. 21, No. 6 (Nov. - Dec., 1973), pp. 1183-1199.*

Greyserman, A., Jones, D. H. & Strawderman, W. E. (2006). Portfolio selection using hierarchical Bayesian analysis and MCMC methods, *Journal of Banking & Finance*, Volume 30, Issue 2, February 2006, pp 669-678.

Griffin, J. E. & Steel, M. F. J. (2006). Inference with non-Gaussian Ornstein–Uhlenbeck processes for stochastic volatility, *Journal of Econometrics*, Volume 134, Issue 2, October 2006, pp 605-644.

Guidolin, M. & Timmermann, A. (2007). Properties of equilibrium asset prices under alternative learning schemes, *Journal of Economic Dynamics and Control*, Volume 31, Issue 1, January 2007, pp 161-217.

Hamilton, J. D. & Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime, *Journal of Econometrics*, Volume 64, Issues 1–2, September–October 1994, pp 307-333.

Kanas, A. (2003). Non-linear Forecasts of Stock Returns, *Journal of Forecasting*, Vol. 22 Issue 4, p 299.

Kılıç, S. B. (2005). Test of the weak form efficient market hypothesis for the Istanbul Stock Exchange by Markov Chains methodology, *University of Social Sciences Of Cukurova,* Vol:14, No:1, pp 333-342.

Lin, S. K., Wang, S. Y. & Tsai, P. L. (2009). Application of hidden Markov switching moving average model in the stock markets: Theory and empirical evidence *International Review of Economics & Finance*, Volume 18, Issue 2, March 2009, pp 306-317.

Liu, H. (2011). Dynamic portfolio choice under ambiguity and regime switching mean returns , *Journal of Economic Dynamics and Control*, Volume 35, Issue 4, April 2011, pp 623-640.

McQueen, G. & Thorley, S. (1991). Are Stock Returns Predictable? A Test Using Markov Chains, *Journal of Finance*, Vol. 46 Issue 1, p239.

Mills, T.C. & Jordanov, J. V. (2003). The size effect and the random walk hypothesis: evidence from the London Stock Exchange using Markov Chains, *Applied Financial Economics*, Vol. 13.

Ryan, T. M. (1973). Security Prices As Markov Processes. *Journal of Financial & Quantitative Analysis*, Vol. 8 Issue 1, pp 17-36.

Tsionas, E. G. (2000). Bayesian model comparison by Markov chain simulation: Illustration using stock market data, *Research in Economics*, Volume 54, Issue 4, December 2000, pp 403-416.

Turner, C. M., Startz, R., & Nelson, C. R. (1989). A Markov Model of Heteroscedastisity Risk and Learning in the Stock market, *Journal of Financial Economics*, 25 (1989) pp 3-22.

Zhang, X. & King, M. L. (2008). Box-Cox stochastic volatility models with heavy-tails and correlated errors, *Journal of Empirical Finance*, Volume 15, Issue3, June2008, pp 549-566.