Application of Hidden Markov Models and Viterbi Algorithm to Characterize the Unemployment in Jordan for the Period (2000-2014)

Dr. Hasan Yasien Tuama
Associate Professor of Applied Statistics
Faculty of Economics and Administrative Sciences,
Zarqa University
Jordan
E-mail: hasantuama@yahoo.com

*** This research is funded by the Deanship of Research in Zarqa University/ Jordan**

ABSTRACT

This study aimed to application of Hidden Markov Models (HMM) and Viterbi algorithm to characterize the unemployment in Jordan for the period (2000-2014). To achieve the study objectives, the study is mainly based on the secondary data related to Unemployment selected from the annual reports of the Jordanian Department of Statistics for the period (2000-2014). The study findings a number of results, including the following:

a. *The unemployment rate system (Y_t) (Observed) is relatively unstable because the variance (fluctuations) ($\sigma^2_1 = 0.976$) between its observations bigger than the variance of the uncertain unemployment system (e^2) (Hidden) ($\sigma^2_0 = 0.326$),*

b. *The unemployment rates estimation in Jordan for the period (2000-2014), seems decreases in the short term.*

Upon the foregoing results, the study reached to a number of conclusions.

KEYWORDS: *Hidden Markov Models, Viterbi Algorithm, Unemployment, Transition Probabilities, Maximum Likelihood Estimation.*

1. INTRODUCTION

The Hidden Markov Model (HMM) is a stochastic model which provides a high level of flexibility for modeling the structure of an observation sequence. It consists of a number of non-observable (Hidden) states and an observable sequence, generated by the individual hidden states (Eyad & Hameed, 2012: 144). Also, the (HMM) is a statistical model which establishes a model for every word through the statistical analysis of large amounts of data with a nite number of states, each associated with a probability distribution. The transitions between states cannot be directly measured (hidden), but in a particular state an observation can be generated. It is the observations and not the states themselves which are visible to an outside observer (Ke & et al., 2008: 305).

The Hidden Markov Model (HMM) is defined as a variant of a finite state machine having a set of hidden states, Q, output observations, O, transition probabilities, A, output probabilities, B, and initial state probabilities, Π. The current state is not observable. Instead of, each state produces an output with a certain probability (B).

Usually the states, Q, and outputs, O, are understood, so an HMM is said to be a triple (A, B, π) (Nikolai, 2010: 1-2).

The study objectives can be summarized as follows:

a. To identify the concept of Markov analysis and Transition Probabilities' matrix.
b. To identify the concept of Hidden Markov Models and viterbi algorithm.
c. Offer some conclusions for the decision makers.

To achieve the study objectives, the researcher formulate the following **hypothesis as null (H_0):**

H_0: There is no statistically significant relationship between the unemployment rates and the uncertain unemployment at the significance level ($\alpha \leq 0.05$).

2. THEORETICAL PART AND LITERATURE REVIEW

2.1. Markov Analysis and Transition Probabilities' Matrix

The Markov analysis is a special case of stochastic or random processes, and look for these operations as a series of situations experienced by the phenomenon through a certain period of time or processes through which a moving object through different periods of time series, called the mentioned operations series by (Markov Chain). In order to identify the Markov processes must be used mathematical analytical methods and conclusions and to clarify the distinctive properties during the development process.

Credited Markov analysis method to the Russian world (A. Markov) (1922- 1952), was limited to use this method in the first instance on the physical applications to study the movement of gas molecules in a closed vessel in order to predict the movement of these molecules in the future.

Based on the foregoing, the Markov analysis defined as "Mathematical and scientific method to analysisof the behavior of different phenomena's during the current period in order to predict the behavior of these phenomena's in the future in any later periods".

An analysis of the Markov method is based on the fundamental assumption that: (any system is dealt with in the first instance be in its initial state, in preparation for the transition to another state), and this assumption based on a certain probability laws called the (Transition Probabilities), which are known as a "transition probabilities of a particular case to another case during a certain period of time".

For example, the probability of transition phenomenon of the case (i) in the current period (n) to another state (j) in the later period (n+1) writes as follows (Touama, 2015: 216):

$$\mathbf{P\{X_{n+1} = j \,/\, X_n = i\} = P_{ij} \quad , \quad \forall\, i , j} \qquad \ldots(1)$$

Whereas:

X_n: Value of the phenomenon in the current period (n).

X_{n+1}: Value of the phenomenon in the subsequent period (n+1).

P_{ij}: Probability of transmission the phenomenon of state (i) to state (j).

The traditional status of the transition probabilities (P_{ij}) values putting in a square matrix $P = [\, P_{ij} \,]$, which takes the following form:

$$\mathbf{P} = \begin{bmatrix} P_{11} & P_{12} & P_{13} & \ldots P_{1n} \\ P_{21} & P_{22} & P_{23} & \ldots P_{2n} \\ & \ldots\ldots & \ldots\ldots & \\ P_{m1} & P_{m2} & P_{m3} & \ldots P_{mn} \end{bmatrix} \qquad \ldots(2)$$

The above matrix called the (Transition Probabilities' Matrix), or sometimes called (Markov Matrix), which represents the matrix of the stochastic or random processes, in which the sum of the probabilities of any row ranks equal to the one, that is:

$$\sum_{j=1}^{n} P_{ij} = 1 \quad , i = 1,2,\ldots,m \qquad \ldots(3)$$

The elements (P_{ij}) of transition probabilities' matrix representing the transition probability from the state (i) to the state (j) by one-step or one time period, if we want to find the probability value of the movement of phenomenon from the state (i) to the state (j) with a limited number of steps or time periods of (m), so the element P_{ij}^{m} can be written as follows:

$$P_{ij}^{m} = P\{ X_{n+m} = j / X_n = i\} \quad , \forall\, i , j \;\; \& \;\; (n, m \in N) \qquad \ldots(4)$$

2.2. Hidden Markov Models

The Hidden Markov models are a statistical tool used for modeling generative sequences characterized by a set of observable sequences. The HMM are widely used in science, engineering and many other areas (speech recognition, optical character recognition, machine translation, bioinformatics, computer vision, finance and economics, and in social science, forecast the weather state, determining the cognitive aspect of the educational process). The HMM framework can be used to model stochastic processes where (Nikolai, 2010: 1-2):

a. The non-observable state of the system is governed by a Markov process.

b. The observable sequences of system have an underlying probabilistic dependence.

The Hidden Markov Model (HMM) is a variant of a finite state machine having a set of hidden states, Q, an output alphabet (observations), O, transition probabilities, A, output (emission) probabilities, B, and initial state probabilities, π. The current state is not observable. Instead, each state produces an output with a certain probability (B). Usually the states, Q, and outputs, O, are understood, so an HMM is said to be a triple, (**A, B,** π).

The parameters set of Hidden Markov Model (HMM) is represented by $\Lambda = (A, B, \pi)$ (Nemeth, 2011: 3):

Whereas:

1.The matrix **A** represent the transition matrix which explain the movement between states, and could be clarified by the following formula:

$$A = \{a_{ij}\} , \text{ Where } \; a_{ij} = P\,[q_{t+1} = S_j \mid q_t = S_i] \;\; , \; 1 \le i, j \le N \;\text{ and } \sum a_{ij} = 1 \qquad \ldots(5)$$

Where:

S: Individual states and are denoted as, $S = \{S_1, S_2, \ldots , S_N\}$.

N: Number of the model states**.**

2. B: represent the observation probability matrix**,** and could be clarified by the following formula:

$$B = \{b_j(k)\} \;, \text{ Where } \; b_j(k) = P\,[O_t = V_k \mid q_t = S_j] \;\; , \; 1 \le k \le M \;, 1 \le j \le N \qquad \ldots(6)$$

Where:

O_t : Observation symbol at time t.

M: Number of distinct observation symbols per state.

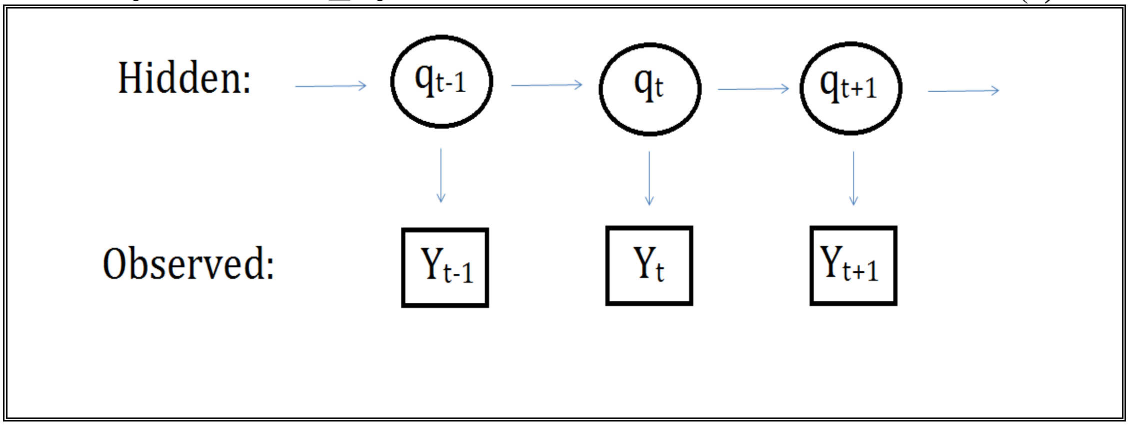
V: Observation symbols and are denoted as $V = \{V_1, V_2, \ldots , V_M\}$.

3. π**:** represent the initial state probabilities, and could be clarified by the following formula:

$$\pi = \{\pi_i\} \;, \text{ Where } \; \pi_i = P\,[q_1 = S_i] \quad , \quad 1 \le i \le N \qquad \ldots(7)$$

Where:

$\pi_i \geq 0$ **and** $\sum \pi_i = 1$ …(8)



Source: Nemeth, Christopher (2011: 2)**.**

Figure 1. Hidden Markov Model with observations Y_t and hidden states q_t

2.3. The Viterbi Algorithm and its assumptions

The Viterbi algorithm chooses the best state sequence that maximizes the likelihood of the state sequence for the given observation sequence. Let $\Phi_t(i)$ be the maximal probability of state sequences of the length **t** that end in state **i** and produce the **t** first observations for the given model (Nikolai, 2010: 4-5):

$$\Phi_t(i) = \max\{P(q(1), q(2), ..., q(t-1) ; O(1), O(2), ... , O(t) \mid q(t) = q_i)\} \quad …(9)$$

As with Viterbi algorithm, we can perform these calculations in the log domain, resulting in the following equation (Eyad & Refeis, 2013: 2519**)** & **(**Nikolai, 2010: 5-7**)**:

$$\text{Ln} \{N(O_t ; \mu_j , \sigma_j)\}= [- (L / 2) \text{Ln} (2\pi) - \sum^{L}_{j=0} \text{Ln} (\sigma_j)] -\sum^{L}_{j=0} (O_t - \mu_t)^2.[1 / 2 \sigma^2_j]$$

And**;** $j = 0,1$ …(10)

The Viterbi Algorithm is defined as a dynamic programming algorithm for finding the most likely sequence of hidden states–called a Veterbi Path – that results in a sequence of observed symbols.

In light of the above, we can explain the **assumptions** of Viterbi Algorithm, and these assumptions are all satisfied in a first-order hidden Markov model**,** as follows:

a. Both of the observed symbols and hidden states must be in a sequence.

b. The two sequences need to be aligned, and an observed symbol needs to correspond to exactly one hidden state

c. Computing the most likely sequence of hidden states (path) up to a certain point **t** must depend only on the observed symbol at point **t**, and the most likely sequence of hidden states (path) up to point (t −1).

To solve the Viterbi algorithm we can use the (Markov's Switching algorithm) for the means and variances for two states (system), as follows (Yousif & Mardan, 2013: 338 -342):

$$y_t = [\mu_0 (1-q_t) + \mu_1 q_t] + [\sigma^2_0 (1-q_t) + \sigma^2_1 q_t] \; \varepsilon_t \quad …(11)$$

Where:

$\varepsilon_t \sim N(0,1)$, q_t take (0 or 1), and y_t represent Unemployment in Jordan.

And the observation (y_t) of the first and second systems in time t, are given through the following formulas:

$$y_t = \{q_t = 0\} = \mu_0 + \sigma^2_0 \varepsilon_t \qquad \ldots(12)$$

$$y_t = \{q_t = 1\} = \mu_1 + \sigma^2_1 \varepsilon_t \qquad \ldots(13)$$

Such that;

Eq.(12): means that the observation y_t according to the first system in time t.

Eq.(13): means that the observation y_t according to the second system in time t.

μ_0, μ_1: Represent the arithmetic means for the first & second systems respectively.

σ^2_0, σ^2_1: Represent the variances for the first & second systems respectively.

And for Markov Switching Model of the first order, and for two hidden states there for the transition probability matrix are given as follows:

$$P = \begin{pmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{pmatrix} \qquad \ldots(14)$$

Where:

$$P_{ij} = Pr\,[q_t = j \mid q_{t-1} = i] \quad , \quad \sum^{L}_{j=0} = P_{ij} = 1 \;, \; \forall i \qquad \ldots(15)$$

$$P(q_t = 0 \mid q_{t-1} = 0) = P_{00} \qquad , \; P_{01} = 1 - P_{00} \qquad \ldots(16)$$

$$P(q_t = 1 \mid q_{t-1} = 1) = P_{11} \qquad , \; P_{10} = 1 - P_{11} \qquad \ldots(17)$$

In light of the above we can get on the solution of the parameters (μ_j, σ_j) and (P_{00}, P_{11}) as follows:

$$\mu_j^{(k)} = \sum^{T}_{t=1} y_t . p(q_t=j \mid y_T; \lambda^{k-1}) / \sum^{T}_{t=1} p(q_t=j \mid y_T; \lambda^{k-1}) \qquad \ldots(18)$$

$$\sigma^{2}_{j}{}^{(k)} = \sum^{T}_{t=1}(y_t - \mu_j^{(k)})((y_t - \mu_j^{(k)}).p\,(q_t= j \mid y_T; \lambda^{k-1}) / \sum^{T}_{t=1} p\,(q_t= j \mid y_T; \lambda^{k-1}) \qquad \ldots(19)$$

$$\sum^{T}_{t=2} p(q_{t=0}, q_{t-1=0} \mid y_T; \lambda^{k-1}) / \sum^{T}_{t=2} p(q_{t=0}, q_{t-1=0} \mid y_T; \lambda^{k-1}) \qquad \ldots(20)$$

$$P_{11}^{(k)} = \sum^{T}_{t=2} p(q_{t=1}, q_{t-1=1} \mid y_T; \lambda^{k-1}) / \sum^{T}_{t=2} p(q_{t=1}, q_{t-1=1} \mid y_T; \lambda^{k-1}) \qquad \ldots(21)$$

2.4. Literature Review

After taking a look at some studies related to Hidden Markov Models, a group of studies relevant to the study theme had been chosen. Atlas & et al., (2000), explained the use of Hidden Markov Models for monitoring machining tool- wear. Boys & et al., (2000), refer to the way of detecting homogeneous segments in DNA sequences by using Hidden Markov Models. Petrushin (2000), discuss the Hidden Markov Models, fundamentals and applications part 2: discrete and continuous Hidden Markov Models. Scott C.S., (2001), explained the importance of the bioinformatics as introduction to Hidden Markov Models, SNUCSE artificial intelligence lab (SCAI). Boys and Henderson (2001), discuss the comparison of reversible jump MCMC algorithms for DNA sequence segmentation by using hidden Markov models. Also, Ramanathan (2006), discuss the applications of Hidden Markov Models. Ramage (2007), explained the Hidden Markov Models fundamentals, CS229 section notes. Yin (2007), indicates to the volatility estimated and price prediction using a Hidden Markov Model with empirical study. Ke & et al., (2008), refers to the HMM speech recognition system based on FPGA. Ropert & et al., (2008), explained the Hidden Markov Models estimation and control. Al-Tuorik, L.S., (2010), evolving the structure of Hidden Markov Model for detection the Micro Aneurysm. Ibrahim, R.M. & et al., (2011), explained the use of Hidden Markov Models in the training and testing of the system by the vocal samples, and the system has been described by blueprints of the standard UML. Nemeth (2011) discuses the outline of the theory behind HMMs, covering areas such as parameter estimation, identification of hidden states and determining the sequence of hidden states . Eyad & Hameed (2012), an efficient face recognition system based on Hidden Markov Model (HMM) and the simplest type "haar" of the discrete wavelet transform (DWT). Abbas, and Farhan (2012) presents an efficient face recognition system based on Hidden Markov Model (HMM) and the simplest type "Haar" of the Discrete Wavelet Transform (DWT). Also, Eyad & Refeis (2013), provides the entrance to study the Influence of different levels of noisy environment on discrimination rate for the speech recognition systems that do not use any type of filters to deal with this problem. Hajo & Florian (2015), explained the Hidden Markov Models with state-dependent mixtures: minimal representation, model testing and applications to clustering.

Some authors as (Zeifman, M.I., and Ingman D., (2003), Sirl D., (2005), Kadim S.K., (2010), Touama H.Y., (2015) indicated to the importance of using Markov analysis, Hidden Markov Models and their applications.

3. THE APPLIED PART

3.1. Collection Data

The study is mainly depend on the secondary data related to Unemployment in Jordan, selected from the annual reports of the Jordanian Department of Statistics. The researcher select the period (2000-2014) in order to achieve the study objectives. As shown in the following Table No. 1:

Table 1. The Unemployment Rates in Jordan for the period (2000-2014)

Years	(1) Unemployment Rates (Y_t)	(2) Estimation Unemployment ($\hat{Y}_t$)	(3) Uncertain Unemployment (e^2) *
2000	13.70	14.50	0.6400
2001	14.70	14.35	0.1225
2002	15.30	14.20	1.2100
2003	14.50	14.05	0.2025
2004	12.50	13.90	1.9600
2005	14.80	13.75	1.1025
2006	14.00	13.60	0.1600
2007	13.10	13.45	0.1225
2008	12.70	13.30	0.3600
2009	12.90	13.16	0.0676
2010	12.50	13.00	0.2500
2011	12.90	12.86	0.0016
2012	12.20	12.71	0.2601
2013	12.60	12.56	0.0016
2014	13.40	12.41	0.9801

(1) Actual values of Unemployment Rates (Y_t) / Jordanian Department of Statistics.
(2) Estimation Unemployment ($\hat{Y}_t$), obtained through applying the estimating equation of Trend line ($\hat{Y}_t = 13.453 - 0.149\ T_i$).
(3) Uncertain Unemployment, obtained through of the square of errors, where ($e = y_{t} - \hat{Y}_t$), and calculated by (SPSS) program.

3.2. RESULTS AND DISCUSSION

3.2.1. Estimation the Unemployment Rates and Uncertain Unemployment

Depending on the relations (18, 19, 20, and 23) we got the Maximum Likelihood Estimates of the Viterbi algorithm by using (MSM) for the Unemployment rates (Y_t)

and the uncertain Unemployment (e^2) by using MATLAB program. As shown in the following Table No. 2:

Table 2. Maximum Likelihood Estimates of the Viterbi algorithm by (MSM*)

Unemployment Rates (Observed) (Y_t)		**Uncertain Unemployment (Hidden) (e^2)**	
μ_1	13.453	μ_0	0.496
σ^2_1	0.976	σ^2_0	0.326.
P_{11}	0.618	P_{00}	0.615

(*) MSM : means Markov Switching Model.

The results listed in table (2) above, can be adopted to distinguish between the unemployment rate (Y_t) and the uncertain unemployment (e^2) according to the values of means and variances, where the results refers to:

a. The unemployment rate system (Observed) (Y_t) is relatively unstable because the variance (fluctuations) (σ^2_1= 0.976) between its observations bigger than the variance of the uncertain unemployment system (e^2) (Hidden) (σ^2_0 = 0.326), and can be considered the uncertain unemployment system is relatively stable system, because the mean of this system which equal to (μ_0 = 0.496) is less than the mean of the unemployment rate (μ_1 = 13.453), and the variance (fluctuation) of the uncertain unemployment less than the variance (fluctuations) for the unemployment rate.

b. The estimates of the transition probabilities (0.618) and (0.615), indicates to fixed of continuity for the two systems and its rapprochement with a very slight differences between them.

3.2.2. Test the Study Hypothesis

H_0: There is no statistically significant relationship between the unemployment rates and the uncertain unemployment at the significance level ($\alpha \leq 0.05$).

To test the previous hypothesis, was used the correlation coefficient (Pearson), as shown in the following Table No. 3:

Table 3. The Correlation coefficients (Pearson) between the unemployment rates and the uncertain unemployment

		Unemployment Rates (Yt)	Uncertain Unemployment (e2)
Unemployment Rates (Yt)	Pearson Correlation	1.000	.189
	Sig. (2-tailed)	.	.500
	N	15	15
Uncertain Unemployment (e2)	Pearson Correlation	.189	1.000
	Sig. (2-tailed)	.500	.
	N	15	15

The results in Table (3), explained that there is no statistically significant relationship between the unemployment rates and the uncertain unemployment at the significance level (α = 0.05). Which is supported by the statistical significant (p-value) for the correlation coefficient, and this value is greater than the significance level (α = 0.05). This means that is **not reject** the null hypothesis (H_0).

In light of the previous results concluded the irrelevance and significance of a causal relationship bidirectional between the unemployment rates and the uncertain unemployment.

4. CONCLUSIONS

This part includes the most important conclusions in light of the results, as follows:

a. The unemployment rate (Observed) (Y_t) is relatively unstable because the variance (fluctuations) ($\sigma^2_1 = 0.976$) between its observations bigger than the variance of the uncertain unemployment (e^2) (Hidden) ($\sigma^2_0 = 0.326$),

b. There is no statistically significant relationship between the unemployment rates and the uncertain unemployment at the significance level ($\alpha = 0.05$).

c. The results of the unemployment rates estimation in Jordan for the period (2000-2014) according to the estimation model of the general Trend ($\hat{Y}_t = 13.453 - 0.149\ T_i$), that the estimates seems decreases in the short term.

REFERENCES

Atlas, L. & et al., (2000), Hidden Markov Models for monitoring machining tool-wear, Submitted to IEEE Symposium.

Al-Tuorik, L.S., (2010), Evolving the Structure of Hidden Markov Models for Micro Aneurysm Detection, Workshop on Computational Intelligence, 8-10/12/2010, Essex, UK.

Boys, R., Henderson, D., and Wilkinson, D., (2000), Detecting homogeneous segments in DNA sequences by using Hidden Markov Models, Journal of the Royal Statistical Society: Applied Statistics, Series C, 49(2): 269-285.

Boys, R. and Henderson, D., (2001), A comparison of reversible jump MCMC algorithms for DNA sequence segmentation using hidden Markov models, Computer. Science and Statistics, 3(3): 35-49.

Eyad, I. A. & Hameed, R. F., (2012), Face Recognition using DWT with HMM, Engineering. & Technical journal, 30(1): 142-154.

Eyad, I.A. & Refeis, A.A., (2013), Influence of Noisy Environment on the Speech Recognition Rate Based on the Altera FPGA, Engineering. & Technical journal, 30(1): 2513-2530.

Hajo, Holzmann & Florian, Schwaiger, (2015), Hidden Markov models with state dependent mixtures: minimal representation, model testing and applications to clustering, Statistics and Computing, 25(6): 1185-1200.

Ibrahim, R.M. & et al., (2011), The use of Hidden Markov Models to identify the judgment Aleqlab, Communications of the Arab Computer Society, 4(2): 1-11.

Jarris, J.P., (2000), Graph-theoretic Analysis of Finite Markov Chains. http//www.oir.memphis

Kadim, S.K., (2010), The use of Markov Chains Absorbance in the Study of the Graduates Numbers in the Education College, Uruk Journal for Scientific Research, Faculty of Science, Department of Mathematics and Computer Applications, 3(3).

Ke, S., Hou, Y., Huang, Z. and Li, H., (2008), A HMM Speech Recognition System Based on FPGA, IEEE, Congress on Image and Signal Processing, Vol. 5, PP. 305-309.

Kijmaka M., (2003), Stochastic Processes with Applications to Finance – Chanpam, New York.

Merlo A., Sericola B., and Marie R., (2000), A New Stable for Computing Steadying State Measures for Markov Chains. http://www.irisa.,fr.htm.

Nemeth, Christopher, (2011), Hidden Markov Models with Applications to DNA Sequence Analysis, STOR-i.

OZ, E., (2010), Algorithms of Hidden Markov Model and a Prediction Method on Product Preferences, Journal of Computing, NY, USA, 2(10): 12-17.

Petrushin, V., (2000), Hidden Markov Models: Fundamentals and Applications Part 2: Discrete and Continuous Hidden Markov Models, (Online) Symposium for Electronics Engineers.

Ramage, Daniel, (2007), Hidden Markov Models Fundamentals, CS229 Section Notes, Available at: http://www.cs.colorado.edu/~martin/slp2.html.

Ramanathan, N., (2006), Applications of Hidden Markov Models, University of Maryland, CMSC 828 J.

Ropert, J.E., Lakhdar, A., & John, B.M., (2008), Hidden Markov Models Estimation and Control, 3rd Edition, Department of Systems engineering, Australian.

Scott C.S., (2001), Bioinformatics Introduction to Hidden Markov Models, SNUCSE Artificial Intelligence Lab (SCAI).

Sirl D., (2005), Markov Chains: An Introduction – Review, University of Queensland Australia, p.1.

Touama H.Y., (2009), Operations Researches: Models and Applications, Dar Safa for Publishing and Distribution, Amman, Jordan.

Touama H.Y., (2015), Application of Markovian Models and Transition Probabilitie's Matrix to Analyze the Workforce Movement in Jordanian Productivity Companies, PARIPEX – Indian Journal of Research, 4(6): 215-218.

Yin, Pei, (2007), Volatility Estimated and Price Prediction using A Hidden Markov Model with Empirical Study, A Thesis Submitted to the Dean of the Graduate School.

Yousif, Munaf & Mardan, M. Askar, (2013), The use of Markov switching model to verify the link between the inflation and uncertain inflation in Iraq for the period (1980-2010), Journal of Economic and Administrative Sciences / Iraq, 19(71): 334-350.

Zeifman, M.I., and Ingman D., (2003), Continuous Markovian Model for Unexpected Shift in SPC, Methodology and Computing in Applied Probability.