

# Testing The Order of Markov Dependence in Error Data of IEEE802.11a Standard

Shaghayegh Kordnoori  
Department of Statistics,  
Science and Research Branch, Islamic Azad University, Tehran, Iran  
shaghayegh.kordnoori@srbiau.ac.ir

Hamidreza Mostafaei  
Department of Statistics,  
Tehran North Branch, Islamic Azad University, Tehran, Iran  
h\_mostafaei@iau-tnb.ac.ir

Mohammad Hassan Behzadi  
Department of Statistics,  
Science and Research Branch, Islamic Azad University, Iran  
behzadi@srbiau.ac.ir

## Abstract

The Markov order is a crucial measure of the memory of a process and its information is essential for appropriate simulation of aspects of the process. In this paper, we suggest a robust and straightforward exact significance test based on generating surrogate data to assess the Markov order of a time series. This method is valid for any sample size and certifies a uniform sampling from the set of sequences that definitely have the  $n$ th order characteristics of the observed data. The Markov property and order of IEEE802.11a errors are investigated using this test.

**Keywords:** Markov order, Whittle's formula, Surrogate data, IEEE 802.11a

## Introduction

An evolution in time of a random phenomenon is a process which can be described by a stochastic process. A stochastic process is a set of random variables indexed by a variable  $t$ , mostly indicating time. Markov chains are a crucial mathematical tool in stochastic processes. A Markov chain satisfies the Markov property, which means that future behavior is independent of past behavior when the present is known. It provides powerful, efficient and flexible means for analyzing the dynamic system properties. The Markov chain is a robust statistical modeling method (Davis, 1973) (Krumbein, W. C., Dacey, M. F., 1969) which can explain complicated behavior of different stochastic systems; Moreover, it has a full-grown mathematical mechanism. Markov chains are applied widely in many different fields such as wireless network (Xianda C., Kyung T. K., Hee Y. Y., 2016) (Hlavičková I, 2015) finance (Huang J., Huang W., Chu P., Lee W., Pai H., Chuang C., Wu Y., 2017) (Mostafaei Hr., Kordnoori Sh., Kordnoori Sh., 2016), biology (Ivan, 2015), computer science (Bolch G., Greiner S., Meer H., Trivedi K., 2006), business (Jónás T., Kalló N., Eszter Tóth Z., 2014) , Chemical engineering (Tamir, 1998), diagnoses of disease (Jackson C. H., Sharples L. D. and Thompson S. G., Duffy S. W., Couto E., 2003), industry (Oviedo-Trespalacios O., baena R. P., Mantilla M., Lacouture C., 2014) ,etc. Many textbooks such as (Feller, 1968) (Howard, 1971) (Cinlar, 1975) provide comprehensive introductions into the elements of stochastic and Markov processes.

The Markov order is a substantial measure of the "memory" of a process, and its information is essential for the accurate simulation of the aspects of the process. There are several Markov chain order estimator recommended in the literature. The Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) are two well-known order estimators according to the Maximum Likelihood (Ding J., Tarokh V., Yang Y., 2017) (Katz, 1981). These methods are only valid in the limit of large samples and their efficiency can not be assured in small sample cases. The success of the AIC is chiefly a result of its better fulfillment when compared with the generally known consistent alternatives. Order estimation strategies include the method of Global Dependency Level (GDL) or relative entropy, the method of (Menendez M., Pardo L., Pardo M., Zografos K., 2011) which employs the  $\phi$ -divergence measures (Pardo, 2006) and the efficient determination criterion (EDC) (Zhao L., Dorea C., Goncalves C., 2001). The (EDC) generalizes the AIC and BIC criteria and yields a class of consistent estimators for the order of a Markov chain with finite state space. The Markov order estimators AIC, BIC and EDC were compared in (Dorea C. C. Y., Angelo P., Gonçalves C., 2015). In (Papapetron M., Kugiumtzis D., 2016) parametric significance tests of Conditional Mutual Information (CMI) was proposed for the order estimation. The relative entropy can be used for estimating an order of the Markov chain (Baigorri A. R., Gonçalves C., Angelo P., 2009). In (Merhav N., Gutman M., Ziv J., 1989) the order of a finite Markov source was estimated based on empirically observed statistics. The order of the chain indicates the number of time steps in the past impacting the probability distribution of the present state, which could be greater than one. It was proved that the order more than one gave the best result for assured conditions. For example in (Deni, S. M., Jemain, A. A., Ibrahim K., 2009) it was showed that the first order of the Markov chain model is the best order for the rainfall occurrence during the northwestern and eastern regions of peninsula, while a higher order is fit for the rainfall occurrence during the northeast monsoon season. Hypotheses related to the proper order Markov chain model can be evaluated statistically applying Chi-Square statistic ( $\chi^2$ ) to contrast expected and observed transition frequencies. The usage of a limiting distribution is imprecise because it is only achieved for infinite data sets and hence it is an inadequate approximation for small data sets. In this case, the method based on p-values developed in this paper used which give the exact test statistics even for small data sets. The exact significance test is employed for any sample size. According to this method, the test statistic distribution is explored by sampling from the set of sequences (named as surrogates) that corresponds exactly to the  $n$ th order characteristics of the observed time series. The surrogate data procedure has some perfect features as follows: One sample is generated per shot, for each sample surrogate as the observed sequence, computation time increases linearly with the length of the sequence, samples are chosen uniformly from the set of all available surrogate sequences and any order can be contained. It is now easy to calculate the p-value of a Markov order null hypothesis exactly for any size data set according to this novel method. The p-value is the probability, assuming the null hypothesis, of the test statistic achieving its observed value or one more extreme. It is not the probability of correctness of the null hypothesis. In the time that a very small p-value leads one to reject the null hypothesis, a large p-value only implies that the data is consistent with the null hypothesis, not that the null hypothesis should be accepted; Moreover, the significance threshold for rejection is perfectly up to the user to determine.

In this paper, we first describe briefly the testing Markov order using the  $\chi^2$  statistics for the large sample limit. Next, the exact significance test is explained according to surrogate data generation and Whittle's formula for any sample size. Finally, this exact significance test is applied for the error data of IEEE802.11a based OFDM system to test the Markov property and find the order of the Markov model.

## The Mathematical Model

### Markov Order Test for Large Sample

Suppose  $P$  be a  $N \times N$  matrix with elements  $\{P_{ij}: i, j = 1, \dots, N\}$  a random process  $(X_0, X_1, \dots)$  with finite state space  $S = \{s_1, s_2, \dots, s_N\}$  is said to be a (homogenous) Markov chain with transition matrix  $P$  if for all  $n, \forall i, j \in \{1, \dots, N\}$  and  $\forall i_0, \dots, i_{n-1} \in \{1, \dots, N\}$  we have Eq. 1:

$$\begin{aligned} P(X_{n+1} = S_j | X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = i) \\ = P(X_{n+1} = S_j | X_n = i) = P_{i,j} \end{aligned} \quad (1)$$

Which implies that the future depends on the past only through the present and not on prior states.

Goodness-of-fit tests play a crucial role in applied and theoretical statistics. They are helpful in evaluating whether a statistical model is consistent with available data. The goodness-of-fit test for the Markov assumption is another pivotal issue. The exact goodness-of-fit tests for first and higher order Markov chains is presented in (Besag J., Mondal D., 2013). Testing the Markov assumption is equivalent to test whether the sojourn time in each state follows an exponential distribution. Some usually applied methods testing whether random variables follow an exponential distribution are according to the empirical distribution function (EDF). Following the theory of (Billingsley, 1961) and the suggestions of (Hoel, 1954), we offer the following method for different types of modeling and testing Markov chains. Suppose  $E_w$  be the expected frequency where  $\sum E_w = N-2$  and  $w$  indicate the set of all frequencies which the expected frequency is greater than zero; Furthermore, let  $O_w \geq 0$  be the corresponding frequency from the observed data. The asymptotic  $\chi^2$  test statistic is defined as Eq.2

$$\chi^2 = \sum_w \frac{(E_w - O_w)^2}{E_w} \quad (2)$$

Given the degrees of freedom  $d$ , The distribution  $f(\chi^2; d)$  is known when  $N \rightarrow \infty$  and the p-value can be achieved by integrating  $f(\chi^2)$  over  $\chi^2 \geq \chi^2_{obs}$ . The degrees of freedom is considered as  $N(N-1)$ .

### Markov Order Test for any Sample Size

The described hypothesis test depends on the  $\chi^2$  distribution allowed in the asymptotic limit of infinite data, therefore it is not exact. Evaluating  $\chi^2$  for all admissible sequences that satisfy the null hypothesis is vital to find the exact distribution for finite data. Suppose a sequence of observations  $\mathbf{x} = \{x_1, \dots, x_N\}$  constitutes a Markov chain of order  $n$  and let  $F_{ij}$  be the number of times the state  $i$  transit to state  $j$  in  $\mathbf{x}$ ; Moreover, suppose

that  $\Gamma(\mathbf{x})$  indicate the set of sequences with the same  $F$  and the same beginning and end states as the observed sequence  $\mathbf{x}$ .

The number of sequences with the word transition count  $F$  and start with state  $k$  and end with state  $l$  is determined by Whittle's formula, Eq.3:

$$N_{kl}(F) = \frac{\prod_i F_i!}{\prod_{i,j} F_{ij}!} C_{kl} \tag{3}$$

Where  $F_i$  is the sum of row  $i$  and  $C_{kl}$  is the  $(k,l)$ th cofactor of the matrix

$$F_{kl}^* = \begin{cases} \delta_{ij}^{-F_{ij}/F_i} & \text{if } F_i > 0 \\ \delta_{ij} & \text{if } F_i = 0 \end{cases} \tag{4}$$

The value of (1) is so large that can not be calculated using fixed accuracy. Therefore, we instead calculate the algorithm of (3) applying a Stirling series for the factorial terms Eq.5:

$$\ln x! \sim x \ln x - x + \frac{1}{2} \ln(2\pi x) + \frac{1}{12x} - \frac{1}{360x^3} + \frac{1}{1260x^5} - \frac{1}{1680x^7} \tag{5}$$

When  $x > 16$ .

The fraction of sequences in  $\Gamma(\mathbf{x})$  that have  $\chi^2$  values greater than or equal to  $\chi^2_{obs}$  yields the p-value. The p-value can be estimated to any desired accuracy provided one has a technique of producing uniform random samples from the set  $\Gamma$  even if  $|\Gamma|$  is too large to enumerate all sequences. The previous techniques for generating samples from  $\Gamma$  are not feasible, specifically for higher order testing. We give a practical method in the next part.

### The Algorithm of Surrogate Generation

The Whittle formula is used for generating a sample subset of  $\Gamma$  in a way that the sample is uniform. Given a transition matrix, a Whittle surrogate is a random sequence which is produced with exactly the same transition count and the beginning and end words are the same as the original sequence. This property assures that the transition probabilities are exactly the same as the original sequence. We consecutively extend a surrogate sequence, starting with a first state until all transitions are utilized. At each step the next state is selected according to the number of remaining sequences calculated by Whittle's formula. Moreover, in order to show the reduced transition count resulting from the selection,  $F$  is updated. The algorithm is certified to result in an acceptable surrogate since states are selected probabilistically weighted by the number of available sequences and the states that result in zero valid sequences are never selected.

Let the sequence  $z = \{z_1, \dots, z_N\}$  be the member of  $\Gamma$  beginning with  $z_1 = k$ , ending with  $z_N = l$ , and having the transition count matrix  $F$ . The nominees for the second element  $z_2$  are the set  $\{w | F_{z_1 w} > 0\}$ . For each nominee  $w$  we calculate  $N_{wl}(F')$ , the number of sequences left. Here  $F'_{ij} = F_{ij} - \delta_{z_1 w}$  is the original transition count matrix minus the candidate transition. A nominee is chosen arbitrarily in proportion to the number of remaining sequences. Therefore

$$P(z_2 = w) = \frac{N_{wl}(F')}{N_{z_1 l}(F)} \tag{6}$$

Once  $z_2$  is selected,  $F$  is reset to the relevant  $F'$  and the process is repeated for  $z_3$  and so on until  $z_{N-1}$  is reached; Moreover, if the beginning and end of the observed sequence are identical symbols ( $x_1 = x_N$ ), then each surrogate should start and end with an identical symbols, but it requires not be the same symbol as observed in the data. The initial state is chosen randomly according to the relative probabilities with Eq.3 for every

alternative of  $z_1 = z_N$ . For the case ( $x_1 \neq x_N$ ) then one must select  $z_1 = x_1$  and  $z_N = x_N$ . The advantages of this technique is its ability of generating long surrogates as the complication increases only linearly with  $N$ .

(Pethel S. D., Hahs D. W., 2014) proved that the surrogate data statistics behave as expected in the asymptotic limit even for a small sample and for this purpose, they computed the size and power of asymptotic and the exact  $\chi^2$  test for different orders of Markov processes using 2500 trials with 2500 surrogates. The p-value which is the fraction of these surrogate sequences that has  $\chi^2 \geq \chi^2_{obs}$  was calculated.

The  $(n+1)$ th order entropy rate  $H_{surg}$  which its computation is easier than the  $\chi^2$  test is an alternative statistic as follow Eq.7.:

$$H(x_{t+1}|x_t, \dots, x_{t-n}) = H(x_{t+1}, x_t, \dots, x_{t-n}) - H(x_t, \dots, x_{t-n}) \quad (7)$$

A sequence of observations form a Markov chain of order  $n$  if

$$P(x_{t+1}|x_t, x_{t-1} \dots) = P(x_{t+1}|x_t, \dots, x_{t-n+1}) \quad (8)$$

or

$$H(x_{t+1}|x_t, x_{t-1} \dots) = H(x_{t+1}|x_t, \dots, x_{t-n+1}) \quad (9)$$

### The Application of the Method

The field of wireless communications is expanding rapidly, recently. Error modeling in communication channels is a favorite approach applied for assessing the channel quality, examining the impact of errors, checking and testing the techniques to enhance the channel performance. Modeling wireless communication errors is substantial for simulation-based performance assessment of network protocols or for utilizing information about these error characteristics within a protocol. In this section, we will use the error data of the IEEE802.11a OFDM system for investigating the Markov property and finding its order by the exact significance test. OFDM is a promising modulation scheme for advanced communications networks. This technology is appropriate for high data rates with sufficient robustness to channel imperfections and frequency selective channels. The IEEE802.11a is an amendment to the 802.11 standard for wireless LANs. It is of the requirement that is mostly known as Wi-Fi which enables us easy Internet connections. The IEEE802.11a standard can produce a high level of performance. The comparison of IEEE802.11a with other IEEE802.11 standards can be found in (Abdelrahman R. B. M., Mustafa A. B. A., Osman A. A., 2015). The specification employs a modulation scheme known as Orthogonal Frequency-division multiplexing (OFDM) which is robust against the effects of multipath propagation. The multicarrier structure, low symbol rate, coding and forward error correction of OFDM make it operable in channel conditions degraded by jamming and fading. It is spectrally efficient and can combat Inter Symbol Interference (ISI) and reduce Inter Carrier Interference (ICI); Moreover, the orthogonality preservation methods in this model are much simpler than CDMA or TDMA. The block diagram for Markov modeling of the generated burst errors from the simulation of an OFDM link is given in Figure 1. The serial to parallel/ parallel to serial, modulation, demodulation, Inverse Fast Fourier Transform and Fast Fourier Transform (IFFT and FFT) are included in OFDM block, Moreover, to circularize the channel effect a redundancy known as a cyclic prefix is considered. The impact of both Multipath-fading and AWGN has been incorporated. The error sequence

as a consequence of imperfections in the transmitter, channel and receiver is achieved by comparing the transmitted and received signal.

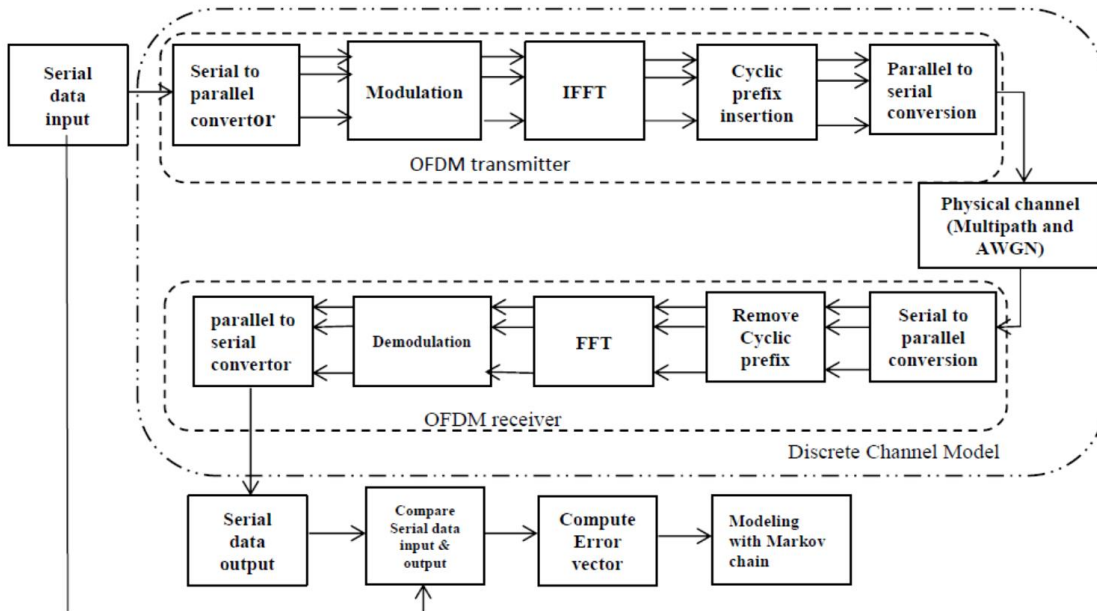


Figure 1: Block diagram for computing error vector and Markov modeling for OFDM.

In our application, the FFT size is 64 and the number of data subcarriers is 52. A  $2 \times 2$  MIMO OFDM system in Rayleigh fading channel with zero force receiver and BPSK modulation is considered. For the 5200 and 7280 binary error data, 400 and 300 surrogate data were produced, respectively to perform the exact significant test. This test is conducted and the p-values were computed for the zero to fourth order Markov test applying the same technique explained above with the block entropy as the discriminating statistic. The alternative hypothesis is one order higher than the order of the null hypothesis. To test the  $n$ th-order null hypothesis with the generated surrogate data, we compare the  $(n+1)$ th order block entropy of the data with the distribution produced by the surrogates. Block entropy histograms of the surrogate data for each order of two examples are shown in Figures 2 and 3, with a solid vertical bar indicating the corresponding block entropy of the error Wi-Fi data. The p-value is shown next to the solid bar which indicates the fraction of surrogate data entropies equal to or less than the entropy of observed data. In both figures, the p-values of order zero admit the existence of the Markov property of Wi-Fi error data at significance level  $\alpha=0.05$ ; Moreover, the p-values of order one ( $p=0.473$  and  $p=0.302$  for 300 and 400 surrogate data, respectively) are consistent with data and certify that Markov model with order one is the best model for modeling the IEEE802.11a error data.

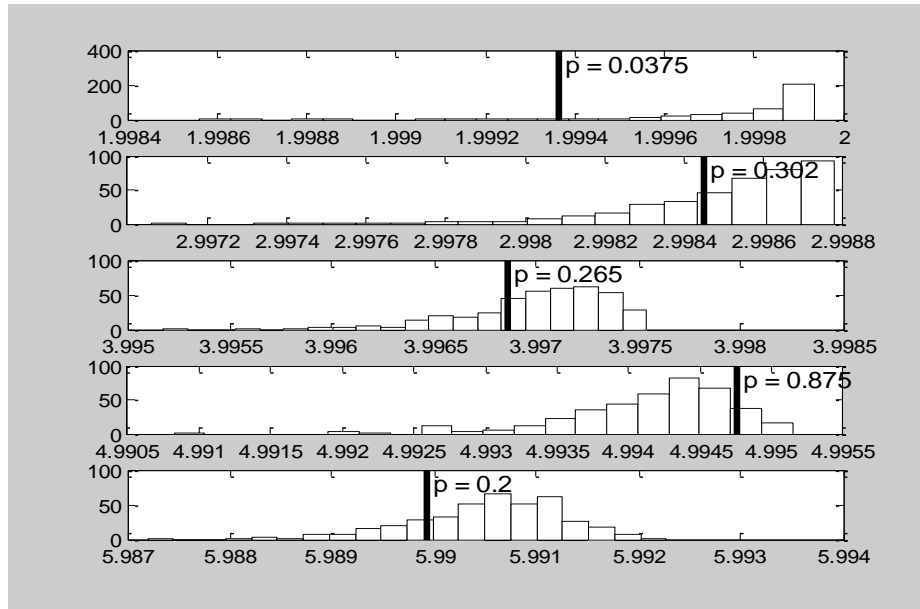


Figure 2: Block entropy histograms of the surrogate data for Markov orders zero through four, respectively. Solid vertical bars show the corresponding block entropy of the errors of the IEEE 802.11a OFDM system with 5200 binary error data and 400 surrogate data.

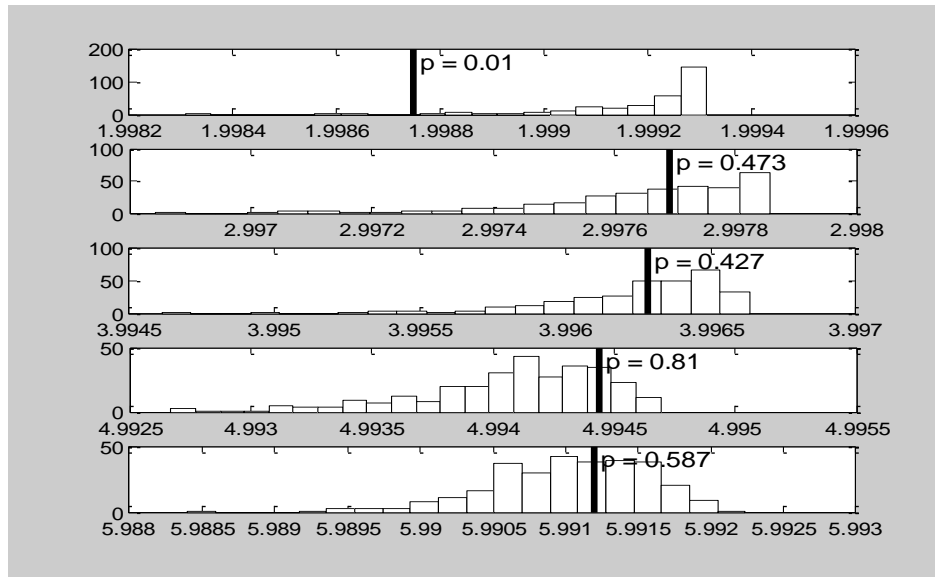


Figure 3: Block entropy histograms of the surrogate data for Markov orders zero through four, respectively. Solid vertical bars show the corresponding block entropy of the errors of the IEEE 802.11a OFDM system with 7280 binary error data and 300 surrogate data.

**Conclusions**

In this paper, we have explained the asymptotic chi square and new exact test of the null hypothesis that a Markov chain is from  $n$ th order. The exact test algorithm is more precise for small data than the asymptotic chi square test. This algorithm is according to Whittle's formula which generates the surrogate data sets with the same word transition

counts as the observed sequence. We have applied the Whittle's algorithm together with the entropy rate statistics for the IEEE802.11a (Wi-Fi) standard errors based on OFDM modulation to test the Markov property and the proper order. We have concluded that these errors have Markov property and their proper order is one. This approach is simple and there is no need for calculating the degrees of freedom or correcting for the small sample size.

## **References**

1. Abdelrahman R. B. M., Mustafa A. B. A., Osman A. A. (2015). A Comparison between IEEE 802.11a, b, g, n and ac Standards. *IOSR Journal of Computer Engineering*, 17 (5, Ver. III), 26-29.
2. Baigorri A. R., Gonçalves C., Angelo P. (2009). Markov Chain Order Estimation and Relative Entropy. Retrieved from <https://arxiv.org>.
3. Besag J., Mondal D. (2013). Exact goodness-of-fit tests for Markov chains. *Biometrics*, 69(2), 488-96.
4. Billingsley, P. (1961). Statistical Methods in Markov Chains. *Annals of Mathematical Statistics*, 32(1), 12-40.
5. Bolch G., Greiner S., Meer H., Trivedi K. (2006). Queueing networks and Markov chains: modeling and performance evaluation with computer science applications. 2nd edition, John Wiley & Sons, Inc., publication.
6. Cinlar, E. (1975). Introduction to stochastic processes. Prentice-Hall, Englewood Cliffs, NJ.
7. Davis, J. C. (1973). Statistics and data analysis in geology. Wiley, New York, 550p.
8. Deni, S. M., Jemain, A. A., Ibrahim K. (2009). Fitting optimum order of Markov chain models for daily rainfall occurrences in Peninsular Malaysia. *Theoretical and Applied Climatology*, 97, 109-121.
9. Ding J., Tarokh V., Yang Y. (2017). Bridging AIC and BIC: A New Criterion for Autoregression. *IEEE TRANSACTIONS ON INFORMATION THEORY*.
10. Dorea C. C. Y., Angelo P., Gonçalves C. (2015). Comparing the markov order estimators AIC, BIC and EDC. *Transactions on Engineering Technologies*, 41-54.
11. Feller, W. (1968). An introduction to probability theory and its applications. vol 1, John Wiley, New York, 3rd Edition.
12. Hlavičková I. (2015). An application of Markov chains in digital communication. *Tatra Mountains Mathematical Publications*, 63(1), 129-137.
13. Hoel, P. (1954). A Test for Markoff Chains (Vols. 41, Parts 3 and 4). Cambridge: University Press.
14. Howard, H. (1971). Dynamic probabilistic systems, Semi Markov and decision processes. John Wiley, New York, vol 2.
15. Huang J., Huang W., Chu P., Lee W., Pai H., Chuang C., Wu Y. (2017). Applying a Markov chain for the stock pricing of a novel forecasting model. *Communications in Statistics-Theory and Methods*, 46(9).
16. Ivan, B. (2015). Markov chain-like quantum biological modeling of mutations. aging and evolution, *Life(Bsel)*, 5(3), 1518-1538.
17. Jackson C. H., Sharples L. D. and Thompson S. G., Duffy S. W., Couto E. (2003). Multistate Markov models for disease progression with classification error, *The Statistician*, 52, part 2, 193-209.



18. Jónás T., Kalló N., Eszter Tóth Z. (2014). Application of Markov Chains for Modeling and Managing Industrial Electronic Repair Processes. *Periodica Polytechnica Social and Management Sciences*, 22(2), 87-98.
19. Katz, R. (1981). On some criteria for estimating the order of a Markov chain. *TECHNOMETRICS*, 23(3).
20. Krumbein, W. C., Dacey, M. F. (1969). Markov chains and embedded Markov chains in geology. *Jour. Internat. Assoc. Math. Geology*, 1(1), 79–96.
21. Menendez M., Pardo L., Pardo M., Zografos K. (2011). Testing the order of Markov dependence in DNA sequences. *Methodology and Computing in applied probability*, 13, 59-74.
22. Merhav N., Gutman M., Ziv J. (1989). On the estimation of the order of a Markov chain and universal data compression. *IEEE Transactions on Information Theory*, 35(5), 1014 – 1019.
23. Mostafaei Hr., Kordnoori Sh., Kordnoori Sh. (2016). Using weighted Markov SCGM(1,1)<sub>c</sub> model to forecast gold/oil, DJIA/gold and USD/XAU ratios. *Malaysian Journal of fundamental and applied sciences*, 12(4), 138-142.
24. Oviedo-Trespalacios O., Baena R. P., Mantilla M., Lacouture C. (2014). The Application of the Markov Chain in Statistical Quality Monitoring. *Int'l Journal of Advances in Mechanical & Automobile Engg. (IJAMAE)*, 1(1), 68-72.
25. Papapetron M., Kugiumtzis D. (2016). Markov chain order estimation with parametric significance tests of conditional mutual information. *Simulation modeling practice and theory*, 61, 1-13.
26. Pardo, L. (2006). *Statistical inference based on divergence measures*. Chapman and Hall, New York.
27. Pethel S. D., Hahs D. W. (2014). Exact significance test for Markov order. *Physica D*, 269, 42-47.
28. Tamir, A. (1998). *Applications of Markov Chains in Chemical Engineering*. 1st Edition, Elsevier Science, 108-119.
29. Xianda C., Kyung T. K., Hee Y. Y. (2016). Integration of Markov random field with Markov chain for efficient event detection using wireless sensor network. *Computer Network*.
30. Zhao L., Dorea C., Goncalves C. (2001). On determination of the order of a Markov chain. *Statistical inference of stochastic processes*, 4, 273-282.