



On the Efficiency of the Maximum Likelihood and Maximum Quasi-Likelihood Estimators in the Second Order Markov Chains

M. Mohammadpour^{†,*}, A. Nematollahi[‡] and M. Yaripour[†]

[†] University of Mazandaran

[‡] Shiraz University

Received: 8/8/2016 Approved: 6/3/2017

Abstract. The present work focuses on the second order Markov chain model which arises in a variety of settings and is well-suited to be modeled in many applications. The efficiency of the maximum quasi-likelihood estimators with the full maximum likelihood estimators for second order Markov chain models are given, besides the limiting normality results on the asymptotic properties of the associated estimates. Some efficiency calculations are also given to discuss the feasibility and computational complexity of the QL approach relative to the full likelihood approach.

Keywords. Maximum likelihood; quasi-likelihood; second order Markov chain; efficiency.

MSC 2010: 62M10; 62M20.

* Corresponding author

1 Introduction

The maximum likelihood (ML) approach is often difficult to carry out numerically, in the theory of many complex models with high dimensional dependencies. The pseudo-likelihood (PL), which is the product over conditional densities, and quasi-likelihood (QL), which is the products of marginal likelihoods for many small subsets, are two different ways to approximate the maximum likelihood estimator. In recent years much progress has been made to use and study the behavior of the maximum pseudo-likelihood and maximum quasi-likelihood estimators. Hu and Zhang (2010) present testing procedures based on pseudo-likelihood ratio statistics with the biased observations and show their asymptotic distributions under the null hypothesis. We refer the reader to Gong and Samaniego (1981); Liang and Self (1996) and Hu et al. (2007), for more details.

The ML method is classic, see Anderson and Goodman (1957) and Billingsley (1961a,b), but the QL results go back to the works of Besag (1974, 1977). The composite likelihood proposed by Lindsay (1988) forms by combining of likelihoods for small subsets of the data, and inherits some good features of the full likelihood which are useful for inference, see Cox and Reid (2004). Quasi-likelihood is a special case of the composite likelihood, in which the quasi-likelihood is formed by the product of the bivariate probability density of all possible pairs of observations. It is attracting researchers in handling high dimensional data analysis when the full likelihood is computationally difficult to evaluate, or intractable when complex dependencies are involved, e.g. in Markovian random fields for spatial and spatio-temporal applications and some type of highly structured hidden Markov models.

The computational feasibility of the pairwise likelihood approach relative to the full likelihood approach are studied by many authors. General discussion on theoretical aspects and possible applications of QL is given in Cox and Reid (2004); Varin (2008) and Varin et al. (2011). Hjort and Varin (2008) discussed and compared three methods of estimation, the ML, a kind of the pseudo-likelihood and the quasi-likelihood, in the case of traditional Markov chain models. In the Markov chain models, the ML would typically be the first choice, because of its simplicity and efficiency. But, it is worth studying the QL alternative in situations where it is still of interest (and appropriate) to compare the statistical behavior of the estimators, see Hjort and Varin (2008). They have also found that the QL strategy is typically preferable to the PL in the Markov chain models.

The present work focuses on the second order Markov chain model which arises in a variety of settings and is well-suited to be modeled in many applications. In chemistry, due to steric effects (arisen from the fact that each atom within a molecule occupies a certain amount of space, Chan and Thouas, 2015), second-order Markov effects may also play a role in the growth of some polymer chains. In the internet applications, Markov models have also been used to analyze web navigation behavior of users. A user's web link transition on a particular web site can be modeled using first- or second-order Markov models and can be used to make predictions regarding future navigation and to personalize the web page for an individual user. In music, a second-order Markov chain can be considered. The first, second, third and fourth order Markov chain can also be used to calculate the transition probability for two-, three-, four- and five-amino-acid sequences, see Wu (1999, 2000). Due to correlation between data in these situations, the exact maximum likelihood estimation, is often a challenge to compute, and thus, we propose the pairwise likelihood approach that aim to balance statistical and computational efficiency, see Varin et al. (2011) for real-world references.

It is well-known that the one dimensional Markov chain of order m is equivalent to an m dimensional first order Markov chain. As a simple example, a second order model for DNA is equivalent to a two dimensional first order model that is comprised of the 16 dinucleotides. A sequence of CGACG (five bases) would be treated as a chain of CG-GA-AC-CG (four states) with each symbol being drawn from the two dimensional setting. Nevertheless, the skeleton of high order model is often preferable for analysis of one dimensional Markov chain of higher order, see Singh (2003). Therefore, in this work, we will work directly with structure of the one dimensional Markov chain of order m , instead of its equivalent m dimensional first order Markov chain.

We concentrate on the QL method and specify its efficiency in the underlying second order Markov chain models. We have argued about the feasibility of the QL approach and discussed its computational complexity relative to competing approaches, however, there is in general, some loss of efficiency compared to ML method. Practical results of the QL and ML are brought into light and some efficiency calculations are presented.

This paper is organized as follows. Preliminaries are given in Section 2. The limiting distribution of ML and QL estimators are established in Section 3. In this section, we also provide some efficiency calculations to demonstrate the feasibility of the proposed QL approach.

2 Preliminaries

Let X_0, X_1, \dots be an irreducible Markov chain of order m (being a positive integer) on a finite state space S , in the sense that

$$P\{X_n = x_n | X_k = x_k, k < n\} = P\{X_n = x_n | X_k = x_k, n - m \leq k < n\}, \quad (1)$$

for all $x_i \in S, i = 0, 1, \dots, n$. Without loss of generality, the elements of S (states) can be considered as a set of integers, $S = \{1, 2, \dots, s\}$.

The model reduces to the classical (or simple) Markov chain, whenever $m = 1$. Throughout what follows, it will be assumed that the Markov chain has stationary transition probability, that is the conditional probabilities in (1) will depend only on the order m , and on the $\{x_k, n - m \leq k < n\}$, and not on the integer n . In the classical Markov chain, the initial probabilities are $p_i = P\{X_0 = i\}, i = 1, 2, \dots, s; \sum_i p_i = 1$ and the transition probabilities are $p_{ij} = \Pr\{X_n = j | X_{n-1} = i\}, i, j = 1, 2, \dots, s; \sum_j p_{ij} = 1$. When the p_{ij} s are specified functions of some unknown parameters θ , the estimation problem of θ by means of a sequence of observations from the chain can be addressed.

In this paper, we treat the case $m = 2$, i.e., the second order Markov chain. The method can be easily extended to the higher order cases in the same manner.

Let X_0, X_1, \dots be a second order Markov chain on a finite state space with stationary transition probabilities. In this case, (1) reduces to

$$\Pr\{X_n = x_n | X_k = x_k, k < n\} = \Pr\{X_n = x_n | X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}\}. \quad (2)$$

On the merits of simplicity, the corresponding parametric model of (2) can be considered of the form

$$p_{a,b,c}(\theta) = \Pr_{\theta}\{X_n = c | X_{n-1} = b, X_{n-2} = a\}, \quad a, b, c = 1, \dots, s, \quad (3)$$

where θ is some underlying parameter vector. The chain is observed from time zero to time N . The aim is to compare the full and pairwise likelihood methods and to see whether there is some loss of efficiency for parameter estimation in the underlying second order Markov chain models, in a large sample setting, i.e., as N grows towards infinity.

3 Main Results

The traditional full maximum likelihood method (ML) maximizes

$$\begin{aligned} l_N(\theta) &= \Pr(X_0 = x_0, X_1 = x_1) \prod_{i=1}^{N-1} \Pr_{\theta}(X_{i+1} = x_{i+1} | X_i = x_i, X_{i-1} = x_{i-1}) \\ &= \prod_{a,b,c} p_{a,b}^{I_{a,b}}(\theta) p_{a,b,c}(\theta)^{N_{a,b,c}}, \end{aligned}$$

where $N_{a,b,c}$ counts the number of transitions from a to b and b to c , i.e.,

$$N_{a,b,c} = \sum_{i=1}^{N-1} I\{(X_{i-1}, X_i, X_{i+1}) = (a, b, c)\},$$

and $I_{a,b} = I(X_0 = a, X_1 = b)$.

We assume that the true model is specified by setting $\theta = \theta_0$. Let $\mathbf{u}_{a,b,c}(\theta)$ (p -vector) and $\mathbf{i}_{a,b,c}(\theta)$ ($p \times p$ matrix) be the first and second derivatives of $\log p_{a,b,c}(\theta)$ and

$$\mathbf{J} = \sum_{a,b,c} p_a p_{a,b} p_{a,b,c} \mathbf{u}_{a,b,c} \mathbf{u}_{a,b,c}^T.$$

Now focus on the ML estimator, say $\hat{\theta}_{ML}$, which under regularity conditions, as stated in Geys et al. (1997); Molenberghs and Verbeke (2005), is the solution of the score function given by

$$\mathbf{U}_N(\theta) = \partial \log l_N(\theta) / \partial \theta = \sum_{a,b,c} N_{a,b,c} \mathbf{u}_{a,b,c} + \sum_{a,b} s I_{a,b} \mathbf{u}_{a,b} = 0.$$

Under the regularity conditions, it is well-known that

$$\sqrt{N}(\hat{\theta}_{ML} - \theta) \rightarrow_d N(\mathbf{0}, \mathbf{J}^{-1}),$$

see e.g. Billingsley (1961). The quasi-likelihood method maximizes the function

$$ql_2(\theta) = \prod_{i=1}^N \Pr_{\theta}(X_{i-1} = x_{i-1}, X_i = x_i) = \prod_{a,b} \{p_a(\theta) p_{a,b}(\theta)\}^{N_{a,b}}, \quad (4)$$

where $\Pr_{\theta}\{X_i = a\} = p_a(\theta)$. The quasi Likelihood (QL) forms by products

of marginal likelihoods for some small subsets. The considered method is a second order QL method. The higher order QL can also be considered. For example, the third order version of QL function can be written as

$$\begin{aligned} ql_3(\theta) &= \prod_{i=1}^{N-1} \Pr_{\theta}(X_{i-1} = x_{i-1}, X_i = x_i, X_{i+1} = x_{i+1}) \\ &= \prod_{a,b,c} \{p_a(\theta)p_{a,b}(\theta)p_{a,b,c}(\theta)\}^{N_{a,b,c}}. \end{aligned} \quad (5)$$

In general, with a fixed order k for QL method,

$$\log ql_k(\theta) = \sum_a N_{a,\dots} \log p_a(\theta) + \sum_{a,b} N_{a,b,\dots} \log p_{a,b}(\theta) + (k-2) \sum_{a,b,c} N_{a,b,c} \log p_{a,b,c}(\theta).$$

In order to establish the limiting distribution of the QL estimator of order k , say $\tilde{\theta}_{QL_k}$, set

$$\mathbf{H} = \sum_a p_a \mathbf{u}_a \mathbf{u}_a^T, \quad \mathbf{F} = \sum_{a,b} p_a p_{a,b} \mathbf{u}_{a,b} \mathbf{u}_{a,b}^T, \quad \mathbf{L} = \sum_{a,b} p_a p_{a,b} \mathbf{u}_{a,b} \kappa_{a,b}^T,$$

and

$$\begin{aligned} \mathbf{G} &= \sum_{a,b} p_a \gamma_{a,b} \mathbf{u}_a \mathbf{u}_b^T, & \mathbf{M} &= \sum_{a,b,c} p_a p_{a,b} p_{a,b,c} \mathbf{u}_{a,b,c} \kappa_{b,c}^T, \\ \mathbf{R} &= \sum_{a,b,c} p_a p_{a,b} p_{a,b,c} \mathbf{u}_{a,b,c} \psi_{b,c}^T, & \mathbf{K} &= \sum_{a,b,c,d} p_a p_{a,b} \omega_{a,b,c,d} \mathbf{u}_{a,b} \mathbf{u}_{c,d}^T, \end{aligned}$$

where

$$\begin{aligned} \kappa_{a,b} &= \sum_{k \geq 0} \sum_d (p_{a,b,d}^{(k)} - p_d) \mathbf{u}_d, & \gamma_{a,b} &= \sum_{k \geq 1} (p_{a,b}^{(k)} - p_b), \\ \psi_{b,c} &= \sum_{k \geq 0} \sum_{e,f} (r_{b,c,e,f}^{(k)} - p_e p_{e,f}) \mathbf{u}_{e,f}, & \omega_{a,b,c,d} &= \sum_{k \geq 1} (r_{a,b,c,d}^{(k)} - p_c p_{c,d}), \\ r_{a,b,c,d}^{(k)} &= \sum_{a_1, \dots, a_{k-1}} p_{a,b,a_1} p_{b,a_1,a_2} p_{a_1,a_2,a_3} \cdots p_{a_{k-2},a_{k-1},c} p_{a_{k-1},c,d}, \\ r_{a,b,c,d}^{(0)} &= p_{a,b,d} \delta_{b,c}, & r_{a,b,c,d}^{(1)} &= p_{a,b,c} p_{b,c,d}, & r_{a,b,c,d}^{(2)} &= \sum_{a_1} p_{a,b,a_1} p_{b,a_1,c} p_{a_1,c,d}, \end{aligned}$$

$p_{a,b,c}^{(k)} = \Pr_{\theta}\{X_{k+1} = c | X_1 = b, X_0 = a\}$ and \mathbf{u}_a denotes the first derivative of $\log p_a(\theta)$.

The following theorem establishes the asymptotic distribution of $\tilde{\theta}_{QL_k}$. The proof is easily derived by using the one-step Taylor analysis of $U_N(\tilde{\theta}_{QL_k}) = \mathbf{0}$, and the result given by Hjort and Varin (2008).

Theorem 1. *Under the regularity conditions,*

$$\sqrt{N}(\tilde{\theta}_{QL_k} - \theta_0) \rightarrow_d N(\mathbf{0}, \mathbf{J}_k^{-1} \mathbf{K}_k \mathbf{J}_k^{-1}),$$

where

$$\begin{aligned} \mathbf{J}_k &= \mathbf{H} + \mathbf{F} + (k-2)\mathbf{J}, \\ \mathbf{K}_k &= \mathbf{H} + \mathbf{G} + \mathbf{G}^T + \mathbf{F} + \mathbf{K} + \mathbf{K}^T + (k-2)^2\mathbf{J} + (\mathbf{L} + \mathbf{L}^T) \\ &\quad + (k-2)(\mathbf{M} + \mathbf{M}^T) + (k-2)(\mathbf{R} + \mathbf{R}^T). \end{aligned}$$

In the following example, we consider a second order Markov chain studied by Azzalini (1983) and discuss about the feasibility and computational complexity of the QL approach relative to the full likelihood approach.

Example 1. Consider the stationary binary second order Markov chain $\{Y_t\}$, with transition probabilities

$$p_{ij} = \Pr(Y_t = 1 | Y_{t-1} = j, Y_{t-2} = i),$$

for $i, j \in \{0, 1\}$. The associated first order Markov chain with states

$$\{(0, 0), (0, 1), (1, 0), (1, 1)\},$$

has transition matrix P and stationary distribution $\pi = (\pi_1, \pi_2, \pi_3, \pi_4)^T$, where

$$\begin{aligned} \pi_2 = \pi_3 &= \{(1 - p_{10})/p_{00} + 2 + p_{01}/(1 - p_{11})\}^{-1}, \\ \pi_1 &= \pi_3(1 - p_{10})/p_{00}, \quad \pi_4 = \pi_3 p_{01}/(1 - p_{11}). \end{aligned}$$

Numerical work has been done using the parametrization

$$p_{ij} = i\theta(1 - \theta) + \frac{1}{2}j(1 - \theta) + \delta, \quad 0 < \theta < 1,$$

where δ is a constant regarded as known, $0 < \delta < \frac{7}{16}$. The behavior of the ML and QL method can be studied and compared by computing the variance of the limiting distribution, algebraically or numerically. The asymptotic variance of the full maximum likelihood estimator is

$$\left\{ \sum_{a,b,c} p_a p_{a,b} p_{a,b,c} \mathbf{u}_{a,b,c} \mathbf{u}_{a,b,c}^T \right\}^{-1},$$

where Azzalini (1983) considered

$$p_0 = \pi_1 + \pi_2, \quad p_1 = \pi_3 + \pi_4,$$

and

$$p_{0,0} = \frac{\pi_1}{\pi_1 + \pi_3}, \quad p_{0,1} = \frac{\pi_3}{\pi_1 + \pi_3},$$

$$p_{1,0} = \frac{\pi_2}{\pi_2 + \pi_4}, \quad p_{1,1} = \frac{\pi_4}{\pi_2 + \pi_4}.$$

Table 1 compares the behavior of ML and QL methods for different order k and $\delta = 0.2$ by their asymptotic variances. As shown in Table 1, when the order k in QL method increases, as we expected, the asymptotic variance of the estimator becomes close to the asymptotic variance of the ML estimator. Table 1 shows that the QL estimators perform well in the considered model. To examine the computation time of the QL and ML estimators, a Monte Carlo simulation is conducted for different sample sizes ($n = 50, 100, 200, 500$). We simulated 10000 samples from the Markov chain for true parameter values $\theta = 0.4$ and $\theta = 0.6$.

Table 1. Asymptotic variances of the estimators for different values of the parameters.

θ	$AVar(\widehat{\theta}_{ML})$	$AVar(\widetilde{\theta}_{QL_2})$	$AVar(\widetilde{\theta}_{QL_3})$	$AVar(\widetilde{\theta}_{QL_{10}})$	$AVar(\widetilde{\theta}_{QL_{100}})$
0.01	0.87	294.29	61.53	3.41	0.88
0.30	4.03	192.86	109.03	16.83	4.02
0.50	2.23	135.01	87.94	17.06	2.34
0.70	1.31	131.79	67.64	8.45	1.35
0.90	0.81	84.40	29.57	2.59	0.81
0.99	0.58	58.25	15.08	1.22	0.57

Table 2. Mean, standard deviation and computational running times of the estimators.

		$\theta = 0.4$						$\theta = 0.6$								
		$\tilde{\theta}_{QL_2}$	$\tilde{\theta}_{QL_3}$	$\tilde{\theta}_{QL_{10}}$	$\tilde{\theta}_{QL_{100}}$	$\hat{\theta}_{ML}$	$\tilde{\theta}_{QL_2}$	$\tilde{\theta}_{QL_3}$	$\tilde{\theta}_{QL_{10}}$	$\tilde{\theta}_{QL_{100}}$	$\hat{\theta}_{ML}$	$\tilde{\theta}_{QL_2}$	$\tilde{\theta}_{QL_3}$	$\tilde{\theta}_{QL_{10}}$	$\tilde{\theta}_{QL_{100}}$	$\hat{\theta}_{ML}$
$n = 50$		0.314(0.024)	0.488(0.022)	0.468(0.009)	0.458(0.008)	0.455(0.005)	0.437(0.027)	0.498(0.010)	0.687(0.009)	0.675(0.006)	0.670(0.006)	0.437(0.027)	0.498(0.010)	0.687(0.009)	0.675(0.006)	0.670(0.006)
	CPU time	0.151	0.162	0.201	0.210	0.250	0.121	0.132	0.166	0.175	0.201	0.121	0.132	0.166	0.175	0.201
$n = 100$		0.493(0.022)	0.472(0.012)	0.467(0.009)	0.459(0.005)	0.452(0.005)	0.432(0.028)	0.493(0.012)	0.683(0.010)	0.673(0.006)	0.678(0.006)	0.432(0.028)	0.493(0.012)	0.683(0.010)	0.673(0.006)	0.678(0.006)
	CPU time	160	0.194	0.213	0.228	0.325	0.167	0.201	0.216	0.306	0.306	0.167	0.201	0.216	0.236	0.306
$n = 200$		0.491(0.021)	0.471(0.012)	0.453(0.008)	0.449(0.004)	0.448(0.004)	0.432(0.057)	0.493(0.012)	0.654(0.008)	0.668(0.005)	0.663(0.005)	0.432(0.057)	0.493(0.012)	0.654(0.008)	0.668(0.005)	0.663(0.005)
	CPU time	0.164	0.272	0.371	0.373	0.500	0.176	0.197	0.198	0.363	0.363	0.176	0.197	0.198	0.220	0.363
$n = 500$		0.492(0.025)	0.460(0.011)	0.450(0.009)	0.441(0.004)	0.440(0.005)	0.432(0.028)	0.493(0.012)	0.652(0.007)	0.664(0.005)	0.660(0.005)	0.432(0.028)	0.493(0.012)	0.652(0.007)	0.664(0.005)	0.660(0.005)
	CPU time	0.236	0.556	0.604	0.645	0.716	0.264	0.589	0.621	0.680	0.671	0.264	0.589	0.621	0.680	0.671

Table 2 provides the mean, standard deviation (in brackets) and the computation time of the estimators. As it was expected, the ML method takes more computational running time than the QL method however the MLEs give the best performance. Also, as the degrees of QL method increases, the computational running times increase. As tedious computations exist in formulating the estimators, more study is needed to confirm the obtained results in the general case.

4 Conclusions

The limiting distribution of ML and QL estimators of the parameters for second order Markov chains are studied. In this paper, the quasi-likelihood (QL), which is the products of marginal likelihoods for many small subsets is applied for the second order Markov chain models. In the Markov chain models, the ML would typically be the first choice, because of its simplicity and efficiency. But it is worth studying the QL alternative in situations where it is still of interest (and appropriate) to compare the statistical behavior of the estimators. Simulation results show that the QL estimators are sufficiently accurate when the order k in the QL estimators increases. It seems that the order k in QL method is a key variable, which should be brought into play. Although more study is needed to confirm the obtained results in the general case.

Acknowledgement

The authors are thankful to the referees for valuable comments.

References

- Anderson, T.W. and Goodman Leo A. (1957). Statistical Inference about Markov Chains. *Ann. Math. Statist.*, **28**, 89-110.
- Azzalini, A. (1983). Maximum Likelihood Estimation of Order m for Stationary Stochastic Processes. *Biometrika*, **70**, 381-387.
- Besag, J.E. (1974). Spatial Interaction and the Statistical Analysis of Lattice Systems. *J. R. Statist. Soc.*, **B 34**, 192-236.
- Besag, J.E. (1977). Efficiency of Pseudolikelihood Estimation for Simple Gaussian Fields. *Biometrika*, **64**, 616-618.

- Billingsley, P. (1961a). Statistical Methods in Markov Chains. *Ann. Math. Statist.*, **32**, 12-40.
- Billingsley, P. (1961b). *Statistical Inference for Markov Processes*. Statistical Research Monographs, Vol. II, University of Chicago Press, Chicago, IL.
- Chen, Q. and Thouas, G. (2015). *Biomaterials: A Basic Introduction*. CRC Press.
- Cox, D.R. and Reid, N. (2004). A Note on Pseudolikelihood Constructed from Marginal Densities. *Biometrika*, **91**, 729-737.
- Hjort, N.L. and Varin, C. (2008). ML, PL, QL in Markov Chain Models. *Scand. J. Statist.*, **35**, 64-82.
- Hu, X.J., Schroeder R.J., Wang, W.C. and Boyett J.M. (2007). Pseudoscore-based Estimation from Biased Observations. *Statist. Med.*, **26**, 2836-2852
- Hu, X.J. and Zhang, Y. (2010). Pseudolikelihood Ratio Test with Biased Observations, *Statist. Papers*, **53**, 387-400.
- Geys, H., Molenberghs, G. and Ryan, L.M. (1997). Pseudo-likelihood Inference for Clustered Binded Data, *Comm. Statist. Theory Methods*, **26**, 2743-2767.
- Gong, G. and Samaniego, F.J. (1981). Pseudo Maximum Likelihood Estimation: Theory and Applications, *Ann. Statist.*, **89**, 861-869.
- Liang, K.Y. and Self, S.G. (1996). On the Asymptotic Behaviour of the Pseudolikelihood Ratio Test Statistic. *J. R. Statist. Soc. B*, **58**, 785-796.
- Lindsay, B.G. (1988). Composite Likelihood Methods. *Contemporary Mathematics*, **80**, 221-239.
- Molenberghs, G. and Verbeke, G. (2005). *Models for Discrete Longitudinal Data*, Springer-Verlag, New York.
- Singh, G.B. (2003). Statistical Modeling of DNA Sequences and Patterns, Chapter 22 in Krawetz, S. A. and Womble, D. D. Introduction to Bioinformatics, Humana Press, Totowa, NJ.
- Varin, C. (2008). On Composite Marginal Likelihoods, *Advances in Statist. Anal.*, **92**, 1-28.
- Varin, C., Reid, N. and Firth, D. (2011). An Overview of Composite Likelihood Methods. *Statist. Sinica*, **21**, 5-42.
- Wu, G. (1999). The First and Second Order Markov Chain Analysis on Amino Acids Sequence of Human Haemoglobin α -Chain and Its Three Variants with Low O₂ Affinity. *Comp. Haematol. Int.*, **9**, 148-151.

Wu, G. (2000). The First, Second, Third and Fourth Order Markov Chain Analysis on Amino Acids Sequence of Human Dopamine β -hydroxylase. *Molecular Psychiatry*, **5**, 448-451.

M. Mohammadpour

Faculty of Mathematical Science,
University of Mazandaran,
Babolsar, Iran.

email: *m.mohammadpour@umz.ac.ir*

A. Nematollahi

College of Science,
Shiraz University,
Shiraz, Iran.

email: *nematollahi@susc.ac.ir*

M. Yaripour

Faculty of Mathematical Science,
University of Mazandaran,
Babolsar, Iran.

email: *jyaripour@yahoo.com*