Empirical likelihood estimation of the Markov-switching model

Paravee Maneejuk¹, Woraphon Yamaka and Songsak Sriboonchitta

Center of Excellence in Econometrics, Faculty of Economics, Chiang Mai University, Chiang Mai, Thailand

¹Email: mparavee@gmail.com

Abstract. The Markov-switching (MS) model is one of the most popular nonlinear time series models in the literature. However, the estimation methods which are normally used to estimate the MS models rely on the assumption of a parametric distribution, which sometimes is considered as a strong assumption. This study, therefore, tries to relax the assumption and develop a more flexible estimator for the MS models that is a maximum empirical likelihood estimation. According to this approach, the parametric likelihood will be replaced by the empirical likelihood function with relatively minor modifications to existing recursive filters. A performance of the suggested estimation method is then evaluated through a Monte Carlo experiment and a real application, the U.S. business cycle. Overall results of both empirical studies indicate that the empirical likelihood could outweigh the classical likelihood estimators.

1. Introduction
A number of studies have shown a nonlinear behavior in economic time series and introduced several models to explain the nonlinear behavior. Among all existing nonlinear models, this study focuses on the Markov-switching model, which is one of the most frequently used nonlinear time series models in the literature, and aims at improving the estimation methods for this model. The original MS-AR model introduced by [1] consists of two states, expansion and contraction, governed by the first order Markov chain. Despite the fact that the model is frequently based on time series data, the classical methods, for example Expected-Maximization, Bayesian, and Maximum Likelihood estimators, are usually used and perhaps the most commonly accepted estimators for parameter estimation. (See, [1], [2] and [3])

However, the model may face to the poor estimation such as bias and inconsistency of the results since the data are assumed to be normally distributed. In the estimation, the time series approach for modeling typically involves a set of strong assumptions. It might not be appropriate to assume that error is a sequence of independent and identically normal distributed random variables [4]. Moreover, [5], [6] and [7] also confirm that the explanatory variable is conditionally non-normal distributed in many applications, thus the assumption of normality then yields biased estimator for estimating parameters of interest. In addition, many researchers concern about the limited data that can lead an underdetermined or ill-posed problem for the observed data, and the conventional estimators are not easily to reach the global optimization. As it is commonly understood, the larger sample size of data can bring the higher probability of finding a significant result [8]. [9] suggests that when the samples are limited, it is often difficult to get a meaningful result.
Thus, the Empirical Likelihood (EL) is suggested additionally to the classical estimation methods to relax a strong assumption of normality and to handle limited data. This approach was first introduced by [10] and considered an alternative to classical likelihood approaches. It has many interesting properties and efficient as parametric likelihood ([11] and [12]). The main idea of EL is to use a maximum entropy discrete distribution supported on the observe data and constrained by nonlinear equations related with the parameters of the model. In short, it is a non-parametric likelihood, which is fundamental for the likelihood-based statistical methodology. A review of EL refers to the study of [13].

In this study, we develop empirical likelihood methods to make inference for parameter estimation under MS-AR model. To the best of our knowledge, the EL has never been used to estimate the MS-AR model. Therefore, our initiation may help to improve the parameter estimation in the MS-AR model and implication. We carry out a Monte Carlo experiment and real application to evaluate a performance of the EL estimator comparing with the conventional estimators.

The rest of this paper is structured as follows. Section 2 provides an explanation on the EL estimator in the MS-AR model. Then, in Section 3, the simulation study is conducted to demonstrate the finite sample performance of the EL approach through Monte Carlo simulations. An empirical comparison of the EL estimator with the conventional maximum likelihood estimator with various distributions is also provided in this section. In Section 4, the MS-AR model with the EL estimator is applied to the U.S. business cycle and Section 5 gives concluding remarks.

2. Methodology

2.1. A Markov-switching autoregressive model

The MS-AR model can be specified as follows:

\[ y_t = \beta_{0,S_t} + \sum_{i=1}^{p} \beta_{i,S_t} y_{t-i} + \varepsilon_{t,S_t}, \quad (1) \]

where \( \varepsilon_{t,S_t} \sim i.i.d. N(0, \sigma^2_{S_t}) \), \( y_t \) is dependent variable and \( y_{t-i} \) is lag dependent variable, in which \( i = 1, ..., p \). \( \beta_{0,S_t} \) denotes regime dependent intercept term and \( \beta_{p,S_t} \) denotes the regime dependent autoregressive coefficients of AR(P). The regime represented by \( S_t \) is considered as an unknown parameter. Let \( \{S_t\}_{t=1}^{h} \) be the finite Markov chain with \( h \) regimes, \( \{1, ..., h\} \). The hidden variable \( S_t \) is generally governed by the first order Markov process, which is defined by

\[ p_{ab} = \Pr(S_{t+1} = b | S_t = a) \quad \text{and} \quad \sum_{b=1}^{M} p_{ab} = 1 \quad ; \quad a, b = 1, ..., M. \quad (2) \]

\( p_{ab} \) is a transitioning probability from regime \( a \) to regime \( b \). All of the transition probabilities can be collected in the transition matrix as follows:

\[ P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1M} \\ p_{21} & p_{22} & \cdots & p_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ p_{M1} & p_{M2} & \cdots & p_{MM} \end{bmatrix}. \quad (3) \]

Usually, the Markov chain is assumed stationary, so the MS-AR process is also stationary, thereby ergodic.

2.2. Constructing empirical likelihood for MS-AR

As this study aims to relax a strong assumption of the parametric likelihood, the distribution of errors of the MS-AR should not be specified. Instead, this study considers the empirical likelihood (EL) as introduced by [14] as an alternative method. So, in this section, we will discuss briefly the concept of empirical likelihood and its relationship with an estimating function. The empirical likelihood function
is the first term of right hand side equation, according to the introduction of [14] and [15], the empirical likelihood is

$$EL(\theta) = \max \prod_{t=1}^{T} p_t.$$  \hspace{1cm} (4) 

By taking logarithm equation (4), the MS-AR problem can take the form as

$$EL(\theta) = \max \sum_{t=1}^{T} \log p_t,$$  \hspace{1cm} (5) 

subject to the constraints:

$$\sum_{t=1}^{T} p_t = 1,$$  \hspace{1cm} (7) 

where $m(y_{t-1}; \theta_s) = \beta_{s,1} \sum_{i=1}^{s} \beta_{s,1} y_{i-1} + \epsilon_{s,1}$ and $\theta_s = \beta_{s,1}, \beta_{s,1}$. These constraints are a vital part of the empirical likelihood estimator as it provides a model information for constructing the empirical likelihood estimates for the MS-AR model. The distribution of the error term is not necessary to be a parametric distribution as the empirical likelihood method allows us to construct a likelihood from observed data. Sometime the high dimensionality of the parameter space $(\theta_s, p_1, ..., p_T)$ makes the above maximization problem difficult to solve and leads to expressions which are hard to maximize. Instead of maximizing $EL(\theta_s)$ with respect to the parameters $(\theta_s, p_1, ..., p_T)$ jointly, the study uses a profile likelihood.

The empirical likelihood, equation (5), is a crucial constrained profile likelihood where the constraints are defined in equation (6) and equation (7), respectively. Then, we can maximize the empirical likelihood at each candidate parameter value $\theta_s$ to obtain the optimal $p_t$. Suppose that we know $\theta_s$, then the empirical likelihood is written as

$$EL(\theta_s, p_1, ..., p_T) = EL(p_1, ..., p_T)$$  \hspace{1cm} (8) 

We maximize this profile empirical likelihood to obtain $(p_1, ..., p_T)$. By conducting the Lagrange multipliers, the study can maximize the empirical likelihood as

$$L(p, \lambda, \lambda') = \sum_{t=1}^{T} \log(p_t) + \lambda_0 \left( \sum_{t=1}^{T} p_t - 1 \right) + \lambda' \sum_{t=1}^{T} p_t \frac{\partial m(y_{t-1}; \theta_s)}{\partial \theta_s} (y_t - m(y_{t-1}; \theta)), \hspace{1cm} (9)$$

where $\lambda \in \mathbb{R}$ is the Lagrange multipliers. It is a straightforward exercise to show that the first order conditions for $L$ with respect to $p_t$, and setting the derivative to zero, the study can find that $\lambda_0 = -T$, and by defining $\lambda = -T \lambda'$, the study obtain the optimal $p_t$ as

$$p_t = \frac{1}{T} \left( 1 + \lambda' \frac{\partial m(y_{t-1}; \theta_s)}{\partial \theta_s} (y_t - m(y_{t-1}; \theta_s)) \right)^{-1}. \hspace{1cm} (10)$$ 

Then, substituting the optimal $p_t$ into the empirical likelihood in equation (4). We then derive the constrained maximization problem form equations (5-7) and obtain the empirical likelihood as follows:

$$EL(\theta_s) = \left( 1 + \lambda' \frac{\partial m(y_{t-1}; \theta_s)}{\partial \theta_s} (y_t - m(y_{t-1}; \theta_s)) \right)^{-1}. \hspace{1cm} (11)$$
Note that the maximum empirical likelihood estimator (MELE) is proposed to estimate the parameters of the model. Without any assumption of that the likelihood distribution, we replace the parametric likelihood with the empirical likelihood and thus the full likelihood of MS-AR is approximated by

$$L(\Theta) = \sum_{S_t=1}^{h} \sum_{r=1}^{T} EL\left(\Theta_{S_t}\right) \cdot \left(E_{S_t} \mid y_r\right),$$  \hspace{1cm} (12)

where \( \left(E_{S_t} \mid y_r\right) \) are the filtered probabilities.

2.3. Hamilton’s filter

In this study, the Hamilton’s filter of [1] is employed to filtering the unobserved variable \( S_t \). As it is unobserved, it is nearly impossible to know which regime prevails at a certain point of time. According to equation (9), the filter probability \( \left(E_{S_t} \mid y_r\right) \) is an important process. We need to filter out the estimated coefficient and variance into different regimes. Hamilton’s filter is determined using the following algorithm.

1. Give values as an initial guess for the transition matrix \( P \).
2. Given past information on \( \Theta_{S_t} \) and \( P \), update the transition probabilities for calculating the likelihood function in each regime at time \( t \). The probability of each regime is updated using equation (12) as shown below:

$$E_{S_t} \mid y_r = \Pr\left(S_t = a \mid \Theta_{S_t}\right) = \frac{EL_a\left(y_r \mid S_t = a, \Theta_{S_t}\right) \Pr\left(S_t = a \mid \Theta_{S_t}\right)}{\sum_{a=1}^{h} EL_a\left(y_r \mid S_t = a, \Theta_{S_t}\right) \Pr\left(S_t = a \mid \Theta_{S_t}\right)}.$$  \hspace{1cm} (13)

Suppose there are \( h \) regimes in the MS-VAR model; \( a=1, \ldots, M \). \( f_a\left(y_r \mid S_t = a, \Theta_{S_t}\right) \) and \( \Pr\left(S_t = a \mid \Theta_{S_t}\right) \) are the likelihood function of regime and the filtered probabilities at time \( t - 1 \), respectively. Then, repeating these two steps for \( t = 1, \ldots, T \).

3. Simulation study

To evaluate the performance of an estimation procedure, we conduct an extensive set of simulations. In the simulation study, we compare the performance of the maximum EL estimator (MELE) with the maximum parametric likelihood estimator. We consider three parametric likelihood: the standard normal distribution (N), student-t distribution (T), skew student-t distribution (ST) as a competing likelihood. The estimation method of three parametric likelihood estimators is described in [16].

In this simulation, we consider two-regime model as a simple example, \( S_t = 1, 2 \). The model can be specified by

$$y_t = \beta_{0,S_t=1} y_{t-1} + \beta_{1,S_t=1} \epsilon_t; \quad y_t = \beta_{0,S_t=2} y_{t-1} + \beta_{1,S_t=2} \epsilon_t.$$  \hspace{1cm} (14)

We set the value for each parameter as follows: \( \sigma_{S_t=1} = 1 \), \( \sigma_{S_t=2} = 0.9 \), \( \beta_{0,S_t=1} = -1 \), \( \beta_{1,S_t=1} = 2 \), \( \beta_{0,S_t=2} = 1.5 \), and \( \beta_{1,S_t=2} = 3 \). Furthermore, the regime variable \( S_t \) is generated from

$$S_t = 1\left(U_t \geq \tau(S_{t-1})\right),$$  \hspace{1cm} (15)

where \( U_t \) is drawn from uniform \([0,1]\) and \( \tau(S_{t-1}) \) is the cumulative probability of transition probability matrix \( P \) in previous regime. Here, we set \( P = \begin{bmatrix} 0.95 & 0.05; 0.95 & 0.05 \end{bmatrix} \).

In the Monte Carlo simulations, the number of observations \( N = 200,400 \) was varied to highlight the effect of sample size on the performance of each model specification. In each experiment, 100
samples were drawn randomly. To make a comparison, we random errors as (1) \(N(0,1)\), (2) \(t(0,1,4)\), (3) \(st(0,1,4,1.5)\) and \(Unif(−1,1)\) where In addition, for the skew and high kurtosis distribution, the degree of freedom for both regimes are set to be \(v_{s,1} = v_{s,2} = 4\), and skew parameters are also set to be \(\gamma_{s,1} = \gamma_{s,2} = 1.5\). Then, we assess the performance of method through the absolute Bias value.

**Table 1. Bias parameters when the error is Normal distribution.**

| Parameter | Empirical N=200 | Normal N=200 | Student-t N=200 | Skewed Student-t N=200 |
|-----------|-----------------|--------------|-----------------|------------------------|
| \(\beta_{0,5,1}\) | 0.3816 | 0.4602 | 0.0858 | 0.0077 |
| \(\beta_{0,5,4}\) | 0.2289 | 0.3943 | 0.1797 | 0.0028 |
| \(\beta_{0,5,2}\) | 0.1949 | 0.1266 | 0.1412 | 0.0626 |
| \(\beta_{0,5,2}\) | 0.2273 | 0.3185 | 0.1151 | 0.0769 |
| \(p_{11}\) | 0.0282 | 0.0241 | 0.0085 | 0.0125 |
| \(p_{22}\) | 0.0114 | 0.0347 | 0.0137 | 0.0109 |

**Table 2. Bias parameters when the error is Student-t distribution**

| Parameter | Empirical N=200 | Normal N=200 | Student-t N=200 | Skewed Student-t N=200 |
|-----------|-----------------|--------------|-----------------|------------------------|
| \(\beta_{0,5,1}\) | 0.7271 | 0.2521 | 0.1470 | 0.0698 |
| \(\beta_{0,5,4}\) | 0.6114 | 0.1668 | 0.0965 | 0.0591 |
| \(\beta_{0,5,2}\) | 0.1630 | 0.0139 | 0.2889 | 0.1015 |
| \(\beta_{0,5,2}\) | 0.1378 | 0.0261 | 0.5139 | 0.4423 |
| \(p_{11}\) | 0.0344 | 0.0494 | 0.1071 | 0.0619 |
| \(p_{22}\) | 0.0406 | 0.0499 | 0.0622 | 0.0573 |

**Table 3. Bias parameters when the error is skewed Student-t distribution.**

| Parameter | Empirical N=20 | Normal N=20 | Student-t N=20 | Skewed Student-t N=20 |
|-----------|----------------|--------------|----------------|------------------------|
| \(\beta_{0,5,1}\) | 0.320 | 0.4152 | 0.1340 | 0.0457 |
| \(\beta_{0,5,4}\) | 0.143 | 0.0157 | 0.5475 | 0.0098 |
| \(\beta_{0,5,2}\) | 2.106 | 0.7473 | 2.6206 | 1.8168 |
| \(\beta_{0,5,2}\) | 1.079 | 0.1612 | 1.7085 | 3.0822 |
| \(p_{11}\) | 0.097 | 0.0327 | 0.0499 | 0.0051 |
| \(p_{22}\) | 0.049 | 0.0456 | 0.9499 | 0.0135 |

**Table 4. Bias parameters when the error is Uniform distribution.**

| Parameter | Empirical N=200 | Normal N=200 | Student-t N=200 | Skewed Student-t N=200 |
|-----------|-----------------|--------------|-----------------|------------------------|
| \(\beta_{0,5,1}\) | 0.0077 | 0.0050 | 1.3960 | 1.1401 |
| \(\beta_{0,5,4}\) | 0.0012 | 0.0066 | 0.6667 | 0.4900 |
| \(\beta_{0,5,2}\) | 1.0416 | 1.0731 | 2.3177 | 2.3615 |
| \(\beta_{0,5,2}\) | 1.1956 | 0.1229 | 1.5693 | 0.7862 |
| \(p_{11}\) | 0.0443 | 0.0450 | 0.0499 | 0.4454 |
| \(p_{22}\) | 0.0352 | 0.0499 | 0.9599 | 0.5298 |
Tables 1-4 report the results of the sampling experiments for the normal, student-t, skewed student-t and uniform error distributions, respectively. In all cases we compute the bias with respect to $\beta_{0,S_1}$, $\beta_{l,S_1}$, $\beta_{0,S_2}$, $\beta_{l,S_2}$, $p_{11}$ and $p_{22}$. In the parametric distribution assumption setting, as expected, we observe the efficiency loss in MELE compared to that parametric MLE. In the other words, the MLE method with the correct model likelihood always had the best performance. However, we can observe that the bias values of the parameters estimated from MELE have an acceptable result as their bias values are close to zero. Moreover, while the error distribution is unknown (uniform distribution), we observe the lowest bias of parameters when using MELE. These experiment study confirms the advantage of the MELE approach, which performs uniformly well over a wide range of error distributions. This estimator shows the appealing efficiency of the MELE when compared to its competitor MLE with normal, student-t and skewed student-t distributions. In addition, it is found that bias is close to zero in some cases. But when the sample size increases, the bias of parameters tends to close to zero in all cases. As a general rule, we can say that the bias tends to approach to zero when the sample size is large, indicating that the estimation technique based on MELE in the MS-AR can provide good asymptotic properties.

4. Empirical study: the U.S. business cycle

The business cycle is one of the most frequently used applications of the Markov-switching (MS) model. A common business cycle shows the rise and fall of the level of production and national output in the economy, usually measured by the gross domestic product (GDP). Although the business cycle can be measured by various techniques, the MS model is often considered one of the most effectively quantitative method for this issue. One reason is that the MS model can have several regimes e.g. expansion and contraction and allows changes in outcomes across the regimes, which make it suitable for illustrating the business cycle. Simultaneously, it provides expected regime durations and amplitudes of the cycle, which are important for generating the stylize facts of business cycle.

Therefore, in this study, we employ the MS model with the suggested estimation method, the MELE, to measure the U.S. business cycle. The quarterly time series of the U.S. GDP from the first quarter of 1947 to the second quarter of 2017 are used in this study. The first part of this section will show the estimated parameters of the model, then the estimated U.S. business cycle will be illustrated through a plot of filtered probabilities. In that picture, the results from the original MS-AR model of [1] will be compared with our suggested one.

| Table 5. Estimated parameters from the MS-AR model form MELE. |
|---------------------------------------------------------------|
| **Regime 1** | **Regime 2** |
| Intercept term | 0.0336*** (0.0477) | 0.0099*** (0.0003) |
| Coefficient | 0.4957*** (0.0083) | −0.2787*** (0.0738) |
| Transition matrix | P1 | P2 |
| P1 | 0.9794 | 0.0206 |
| P2 | 0.0100 | 0.9900 |
| Duration | 48.543 | 100 |

Note: "***" denotes the significance at 1% level and inside (.) is standard error.

First of all, table 5 shows the estimated results of the MS-AR model using MELE. We assume the first-order autoregressive model and a common business cycle with two regimes: expansion and contraction. The estimated result shows a significant regime shift in a stochastic process of the U.S. GDP growth and different estimated parameters across regimes. A coefficient in regime 2 has a negative sign, meaning that growing economy in a previous time can causes a fall in GDP growth at the present time. In addition, intercept term of regime 2 is found less than that of regime 1. These findings suggest that regime 2 is likely to conform to the contraction, whilst regime 1 corresponds to expansion.

The significant regime shift found here may not surprise as the upward and downward trends usually occur in every economy due to changes in economic activities. However, the MS model
provides a transition matrix which allows us to measure the asymmetry of the business cycle in terms of durations and transition probabilities. Firstly, the probability of transitioning from regime 1 (expansion) to regime 2 (contraction) is 0.0206, whilst the chance of remaining in the expansion is 0.9794. Contraction, on the other hand, has the probability 0.99, whilst the chance of switching from the contraction to the expansion is 0.01.

Additionally, the estimated duration shows that U.S. economic contraction last longer than the expansion between 1947 and 2017. The business cycle is approximately 49 quarters in expansion and 100 quarters in contraction. Figure 1 illustrates the business cycle of the U.S. through filtered probabilities. The probabilities lying between 0 and 1 correspond to a chance of the U.S. economy remaining expansion. As shown in figure 1, there are two lines: dash line and solid line. The dash line presents probabilities estimated by maximum empirical likelihood estimator (MELE) and the solid line is estimated by the classical maximum likelihood estimator (MLE). Both estimations just show a relevant result at least in terms of the beginning of the long expansion. The solid line (from classical MLE) shows that the period of expansion begins in 1983 and continues until 2007-08 when the contraction period known as the great recession occurred. However, after the great recession, the U.S. economy recovered and was able to move back to the expansion in the third quarter of 2009.

Even though the estimated result from classical MLE (solid line) conforms to the economic situation at that time, the MELE just provides even more consistent result. For example, the longest expansion may not be as long as the MLE has predicted. According to the National Bureau of Economic Research, the longest expansionary period of the U.S. business cycle starts from the second quarter of 1991 to 2001, and the MELE can capture precisely this golden age of expansion, whilst the MLE is unable to distinguish this period from the whole. On the other hand, the MLE shows that before the long expansion begins, the U.S. economy is almost completely in the contraction except for a period between 1960 and 1965, when the economy could recovery. However, contraction usually occurs more frequently and much shorter than that. For example, it was the 1957 recession due to the Fed’s contractionary policy and lasted around 8 months before recovered. Later, the 1973-75 recession resulting from the OPEC oil embargo occurred and lasted 16 months. These recessions, for instance, can be observed from the estimated result of MELE. According to a theory of business cycle, the contraction is usually followed by a period of recovery and expansion respectively, meaning that there is always a period before the next recession begins. Therefore, the result obtained from MELE is relatively reliable than that of the MLE, based on this dataset, as it can capture other periods of recessions in the history and distinguish one recession from the others, whilst the MLE combines them into a long period of recession.

![Figure 1. The US business cycle (Expansion).](image)
practice as the distribution information of observed data is unknown, thereby inappropriate likelihood function. This study, therefore, considers maximum empirical likelihood estimator (MELE) as an alternative method. It is relatively flexible as it does not assume a certain parametric form for the error distribution, whilst the classical methods do. To show how this technique would work, we conduct a simulation study comparing a performance of the MELE with the MLE. Then, we apply the MS-AR model to the U.S. business cycle and estimate the unknown parameters in the model using MLEL. In addition, the business cycle estimated by MELE is then compared with the cycle obtained from the MLE. The results of the simulation study reveal that the MELE performs uniformly well over a wide range of error distributions. It shows the appealing efficiency of the MELE when compare to the conventional MLE with normal, student-t and skewed student-t distributions. In addition, overall results of the empirical study also suggest in the same way. It is found that the MELE has better performance than the MLE as it can capture precisely a period of recessions and expansions which are more relevant to the real economic situations than that of the MLE.

References
[1] Hamilton J D 1989 A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle *Econometrica* 57(2) 357-384
[2] Kim C J and Nelson C R 1999 Has the US economy become more stable? A Bayesian approach based on a Markov-switching model of the business cycle *Review of Economics and Statistics* 81(4) 608-616
[3] Douc R, Moulines E and Ryden T 2004 Asymptotic properties of the maximum likelihood estimator in autoregressive models with Markov regime *The Annals of Statistics* 32(5) 2254-2304
[4] Howrey 1980 The Role of Time Series Analysis in Econometric Model Evaluation *The National Bureau of Economic Research*
[5] Lourens S, Zhang Y, Long J D and Paulsen J S 2013 Bias in estimation of a mixture of normal distributions *Journal of biometrics & biostatistics* 4
[6] Sriboonchitta S, Yamaka W, Maneejuk P and Pastipatkul P 2017 A Generalized Information Theoretical Approach to Non-linear Time Series Model *In Robustness in Econometrics* (Springer International Publishing) 333-348
[7] Maneejuk P, Yamaka W and Sriboonchitta S 2018 A Markov-Switching Model with Mixture Distribution Regimes *In International Symposium on Integrated Uncertainty in Knowledge Modelling and Decision Making* (Springer Cham) 312-323
[8] Peto R, Pike M C, Armitage P, Breslow N E, Cox D R, Howard S V and Smith P G 1977 Design and analysis of randomized clinical trials requiring prolonged observation of each patient II Analysis and examples *British journal of cancer* 35(1) 1
[9] Button K S, Ioannidis J P, Mokrysz C, Nosek B A, Flint J, Robinson E S and Munafò M R 2013 Power failure: why small sample size undermines the reliability of neuroscience *Nature Reviews Neuroscience* 14(5) 365-376
[10] Owen A B 1998 Empirical Likelihood John Wiley & Sons Inc
[11] Mykland P A 1999 Bartlett identities and large deviations in likelihood theory *Annals of statistics* 1105-1117
[12] Varyiath A M, Chen J and Abraham B 2010 Empirical likelihood based variable selection *Journal of Statistical Planning and Inference* 140(4) 971-981
[13] Chen S X and Van Keilegom I 2009 A review on empirical likelihood methods for regression *Test* 18(3) 415-447
[14] Owen 1991 Empirical Likelihood for Linear Models *The Annals of Statistics* 19(4) 1725 – 1747
[15] Owen 2001 Empirical Likelihood (New York: Chapman and Hall/CRC)
[16] Maneejuk P, Yamaka W and Sriboonchitta S 2018 A Markov-Switching Model with Mixture Distribution Regimes *In International Symposium on Integrated Uncertainty in Knowledge Modelling and Decision Making* (Springer Cham) 312-323

8