Chilean Journal of Statistics
Vol. 7, No. 1, April 2016, 55–65

<u>Time Series</u>
Research Paper

Bernoulli Difference Time Series Models

Abdulhamid A. Alzaid, Maha A. Omair* and Weaam. M. Alhadlaq

Department of Statistics and Operations Research, King Saud University, Riyadh, Kingdom of Saudi Arabia

(Received: 24 May 2015 · Accepted in final form: 18 April 2016)

Abstract

In this paper, we present a Bernoulli difference Markov model and a Bernoulli difference time series model based on Jacobs-Lewis mixture method. The limiting distribution of Jacobs-Lewis mixture model is obtained. The Bernoulli difference Markov model allows for positive and negative correlation. Maximum likelihood, conditional maximum likelihood, and Yule Walker methods of estimation are considered. Simulations are carried out. The paper concludes with an analysis of a real data set.

Keywords: Bernoulli difference distribution · Jacobs-Lewis model · Markov model · Stock Market · Time series model.

Mathematics Subject Classification: Primary 62M05, 62M10 · Secondary 60E05, 46N30.

1. Introduction

In recent years, increasing attention has been given to the analysis of discrete variate time series, reflecting a need for models that accounts for the count nature of the data. Discrete time series of small counts occur in many areas; some examples are the number of customers waiting to be served at a counter recorded at discrete points in time, the monthly number of accidents in a manufacturing plant, the monthly number of AIDS diagnosed people in a specific area, daily count of epileptic seizures of a patient.

Models for non-negative discrete time series have been suggested by many researchers. Jacobs and Lewis (1978 a, b, 1983) defined discrete autoregressive-moving average (DARMA) models based on mixtures. Al-Osh and Alzaid (1987), Alzaid and Al-Osh (1990) and McKenzie (1986) proposed the integer-valued autoregressive-moving average (INARMA) process based on thinning operator.

In many applications of modeling count series, we are faced with non-stationary process. This necessitates taking the difference which gives rise to the need of time series models taking negative and positive integers. Kim and Park (2008) defined the signed binomial thinning operator and proposed a non-stationary integer valued autoregressive model. Karlis and Anderson (2009) introduced a time series process on Z: a ZINAR model.

*Corresponding author. Email:Maomair@ksu.edu.sa

ISSN: 0718-7912 (print)/ISSN: 0718-7920 (online)

http://www.soche.cl/chjs

Zhang et al. (2010) considered inference of integer valued autoregressive models (INAR(p)) with signed generalized power series thinning operator. Kachour and Truquet (2011) defined a p-order signed integer-valued autoregressive model (SINAR(1)). Bulla et al. (2012) introduced a bivariate first order signed integer valued autoregressive process. Alzaid and Omair (2014) introduced a Poisson difference integer-valued autoregressive model of order one.

Odhah (2013) studied the Bernoulli difference distribution in details and discussed some of its properties. Omair et al. (2016) defined a trinomial difference distribution and applied the distribution to traffic accident.

The following gives the definition and the moments of the distribution.

Definition 1.1 A random variable X has Bernoulli difference distribution with parameters α and β denoted by $X \backsim BerD(\alpha,\beta)$ if it takes the values -1, 0, and 1 with the probabilities β, $1-\alpha-\beta$, and α respectively. The probability mass function of $BerD(\alpha,\beta)$ can be written as

$$f(x;\alpha,\beta)=\alpha^{\frac{|x|+x}{2}}\beta^{\frac{|x|-x}{2}}(1-\alpha-\beta)^{1-|x|};x=-1,0,1,\text{where } \alpha,\beta\geq 0 \text{ and } \alpha+\beta\leq 1.$$

The mean and the variance are

$$E(X)=\alpha-\beta,\ V(X)=\alpha(1-\alpha)+\beta(1-\beta)+2\alpha\beta.$$

The aim of this paper is to introduce two time series models with first order autoregressive structure for Bernoulli difference distribution. The first model is the Markov model which assumes that the conditional distribution of $Y_t|Y_{t-1}$ follows Bernoulli difference distribution. This model allows for positive and negative correlation. The second model is the Jacobs-Lewis mixture model which is based on the Jacobs-Lewis mixture methods with Bernoulli difference distribution. This model can only be fitted to positively correlated data.

The remainder of this paper proceeds as follows: In Section 2, Bernoulli difference Markov model is introduced and estimation of the parameters were discussed. In Section 3, Jacobs-Lewis Mixture Model is presented and the estimates of the parameters were obtained. In Section 4, a simulation studies are performed to enhance the conclusions for the two models. In Section 5, a real application of the Bernoulli Difference (BerD) Markov model is illustrated. Finally, we present our conclusions in Section 6.

2. Bernoulli Difference Markov Model

We consider here a Markov chain model with marginal Bernoulli difference distribution. The model is defined as follow.

Let $\{Y_t : t = 1,2,...\}$ be a Markov sequence of Bernoulli difference random variables such that

$$\alpha(y_{t-1})=P(Y_t=1|Y_{t-1}=y_{t-1})=\frac{e^{\delta_{00}+\delta_{01}y_{t-1}}}{1+e^{\delta_{00}+\delta_{01}y_{t-1}}+e^{\delta_{10}+\delta_{11}y_{t-1}}}$$

and

$$\beta(y_{t-1})=P(Y_t=-1|Y_{t-1}=y_{t-1})=\frac{e^{\delta_{10}+\delta_{11}y_{t-1}}}{1+e^{\delta_{00}+\delta_{01}y_{t-1}}+e^{\delta_{10}+\delta_{11}y_{t-1}}}$$

This implies that the conditional distribution of Y_t given $Y_{t-1} = y_{t-1}$ is

$BerD(\alpha(y_{t-1}), \beta(y_{t-1}))$. Hence, the conditional mean is

$$E(Y_t|Y_{t-1}) = \alpha(y_{t-1}) - \beta(y_{t-1}),$$

the conditional variance is

$$Var(Y_t|Y_{t-1}) = \alpha(y_{t-1}) + \beta(y_{t-1}) - [\alpha(y_{t-1}) - \beta(y_{t-1})]^2$$

and the covariance between any two adjacent variables is

$$\begin{aligned} Cov(Y_t, Y_{t-1}) &= E[Cov(Y_t, Y_{t-1}|Y_{t-1})] + Cov[E(Y_t|Y_{t-1}), E(Y_{t-1}|Y_{t-1})] \\ &= Cov[\alpha(Y_{t-1}) - \beta(Y_{t-1}), Y_{t-1}]. \end{aligned}$$

The conditional maximum likelihood function is

$$CL = f(y_1, y_2, ..., y_t|y_1) = \prod_{t=2}^{n} f(y_t|y_{t-1}) = \prod_{t=2}^{n} \alpha_t^{\frac{|y_t|+y_t}{2}} \beta_t^{\frac{|y_t|-y_t}{2}} (1 - \alpha_t - \beta_t)^{1-|y_t|}.$$

The conditional maximum likelihood estimators $\hat{\delta}_{00}, \hat{\delta}_{01}, \hat{\delta}_{10}$ and $\hat{\delta}_{11}$ are obtained by solving the following four nonlinear equations

$$\begin{aligned} \frac{\partial logCL}{\partial \delta_{00}} &= \sum_{t=2}^{n} \frac{|y_t| + y_t}{2} - \sum_{t=2}^{n} \frac{e^{\delta_{00}+\delta_{01}y_{t-1}}}{1 + e^{\delta_{00}+\delta_{01}y_{t-1}} + e^{\delta_{10}+\delta_{11}y_{t-1}}} = 0, \\ \frac{\partial logCL}{\partial \delta_{01}} &= \sum_{t=2}^{n} \frac{|y_t| + y_t}{2} y_{t-1} - \sum_{t=2}^{n} \frac{e^{\delta_{00}+\delta_{01}y_{t-1}}}{1 + e^{\delta_{00}+\delta_{01}y_{t-1}} + e^{\delta_{10}+\delta_{11}y_{t-1}}} y_{t-1} = 0, \\ \frac{\partial logCL}{\partial \delta_{10}} &= \sum_{t=2}^{n} \frac{|y_t| - y_t}{2} - \sum_{t=2}^{n} \frac{e^{\delta_{10}+\delta_{11}y_{t-1}}}{1 + e^{\delta_{00}+\delta_{01}y_{t-1}} + e^{\delta_{10}+\delta_{11}y_{t-1}}} = 0, \text{ and} \\ \frac{\partial logCL}{\partial \delta_{11}} &= \sum_{t=2}^{n} \frac{|y_t| - y_t}{2} y_{t-1} - \sum_{t=2}^{n} \frac{e^{\delta_{10}+\delta_{11}y_{t-1}}}{1 + e^{\delta_{00}+\delta_{01}y_{t-1}} + e^{\delta_{10}+\delta_{11}y_{t-1}}} y_{t-1} = 0. \end{aligned}$$

The maximum likelihood estimator $\hat{\boldsymbol{\delta}} = (\hat{\delta}_{00}, \hat{\delta}_{01}, \hat{\delta}_{10}, \hat{\delta}_{11})$ is asymptotically normally distributed $N_4(\boldsymbol{\delta}, I^{-1}(\boldsymbol{\delta}))$ or more accurately $\sqrt{n}(\hat{\boldsymbol{\delta}} - \boldsymbol{\delta})$ is asymptotically $N_4(0, nI^{-1}(\boldsymbol{\delta}))$ where $I(\boldsymbol{\delta})$ is the Fisher information matrix with entries $I_{ij,kl} = E(-\frac{\partial^2 logCL}{\partial \delta_{ij} \partial \delta_{kl}})$.

If $|\delta_{ij}|$ are finite then clearly from the definition of the model all the probabilities of the transition matrix are positive. Hence, all the states are communicates. Therefore, the stationary distribution always exists and can be obtained as the limit of the power of transition matrix.

3. Jacobs-Lewis Mixture Model

Jacobs and Lewis (1978 a, b) introduced a model for obtaining a sequence of discrete random variables with first order Markov dependence and with a given marginal distribution. The process is a Markov chain and has correlation structure of a first order autoregressive process (AR(1)). We define Jacobs-Lewis Bernoulli difference AR(1) model as

$$Y_t = U_t Y_{t-1} + (1 - U_t)\varepsilon_t, \quad t = 2, \ldots, n,$$

where U_t are independent identically distributed (i.i.d.) Ber(p) and independent from $\varepsilon_t \backsim BerD(\alpha, \beta)$, also ε_2 is independent of Y_1.

The conditional probability of Y_t given Y_{t-1} is

$$f(y_t|y_{t-1}) = p\delta_{y_t}(y_{t-1}) + (1-p)\alpha^{\frac{|y_t|+y_t}{2}}\beta^{\frac{|y_t|-y_t}{2}}(1-\alpha-\beta)^{1-|y_t|},$$

where $\delta_{y_t}(y_{t-1}) = \begin{cases} 1, & \text{if } \ y_t = y_{t-1} \\ 0, & \text{Otherwise.} \end{cases}$

Proposition 3.1

(1) The mean is $E(Y_t) = p^t E(Y_1) + (1-p^t)(\alpha - \beta)$.
(2) The variance is $V(Y_t) = p^t E(Y_1^2) + (1-p^t)(\alpha+\beta) - (\alpha-\beta)^2$.
(3) The covariance between two successive observations is

$$Cov(Y_t, Y_{t-1}) = pVar(Y_{t-1}) = p^{t+1}E(Y_1^2) + p(1-p^t)(\alpha+\beta) - (\alpha-\beta)^2.$$

Thus,

$$Cov(Y_t, Y_{t-k}) = p^k Var(Y_{t-k}) = p^{t+k}E(Y_1^2) + p^k(1-p^t)(\alpha+\beta) - (\alpha-\beta)^2.$$

Now if we assume that the initial distribution of Y_1 is $BerD(\alpha, \beta)$, then $\{Y_t\}$ is a Bernoulli difference stationary process.

Corollary 3.2

(1) $\lim E(Y_t) = \alpha - \beta$.
(2) $\lim V(Y_t) = \alpha + \beta - (\alpha-\beta)^2$.
(3) $\rho_k = corr(Y_t, Y_{t-k}) = p^k$.

Note that this model can only be fitted to positive correlated data, contrary to the Bernoulli difference Markov model which can be fitted to both positively and negatively correlated. Proof of proposition 3.1 is provided in Appendix A.

3.1 Estimation of the Parameters

Let $\{Y_t; t = 1, 2, ..., n\}$ be a Jacobs-Lewis Bernoulli difference AR(1) process. In this section, three methods of estimation are considered: Yule worker method, conditional maximum likelihood method and maximum likelihood method.

The Yule-Walker estimators $\hat{\alpha}_{YW}, \hat{\beta}_{YW}$ and $\hat{p}_{YW}$ of the parameters α, β and p respectively are obtained from the following moments equations

$$\begin{aligned} \bar{Y} &= \hat{\alpha}_{YW} - \hat{\beta}_{YW}, \\ s_y^2 &= \hat{\alpha}_{YW} + \hat{\beta}_{YW} - (\hat{\alpha}_{YW} - \hat{\beta}_{YW})^2, \text{ and} \\ r &= \hat{p}_{YW}, \end{aligned}$$

where $\bar{Y}$, s_y^2 and r are the sample mean, the sample variance and the sample autocorrelation respectively.

Therefore; $\hat{\alpha}_{YW} = \frac{s_y^2+\bar{Y}^2+\bar{Y}}{2}$, $\hat{\beta}_{YW} = \frac{s_y^2+\bar{Y}^2-\bar{Y}}{2}$ and $\hat{p}_{YW} = r = \frac{\sum_{t=2}^{n}(Y_t-\bar{Y})(Y_{t-1}-\bar{Y})}{\sum_{t=1}^{n}(Y_t-\bar{Y})^2}$.

The conditional likelihood function is

$$CL = \prod_{t=2}^{n} f(y_t|y_{t-1}) = \prod_{t=2}^{n}[p\delta_{y_t}(y_{t-1}) + (1-p)\alpha^{\frac{|y_t|+y_t}{2}}\beta^{\frac{|y_t|-y_t}{2}}(1-\alpha-\beta)^{1-|y_t|}]$$

The conditional maximum likelihood estimators $\hat{\alpha}_{CML}$, $\hat{\beta}_{CML}$ and $\hat{p}_{CML}$ are obtained by solving the following nonlinear equations

$$\begin{aligned}
\frac{\partial logCL}{\partial\alpha} &= \sum_{t=2}^{n}\frac{1}{A}(1-p)\alpha^{\frac{|y_t|+y_t}{2}-1}\beta^{\frac{|y_t|-y_t}{2}}(1-\alpha-\beta)^{-|y_t|}[\frac{|y_t|+y_t}{2}(1-\beta)+(\frac{|y_t|-y_t}{2}-1)\alpha] \\
&= 0, \\
\frac{\partial logCL}{\partial\beta} &= \sum_{t=2}^{n}\frac{1}{A}(1-p)\alpha^{\frac{|y_t|+y_t}{2}}\beta^{\frac{|y_t|-y_t}{2}-1}(1-\alpha-\beta)^{-|y_t|}[\frac{|y_t|-y_t}{2}(1-\alpha)+(\frac{|y_t|+y_t}{2}-1)\beta] \\
&= 0, \text{ and} \\
\frac{\partial logCL}{\partial p} &= \sum_{t=2}^{n}\frac{1}{A}[\delta_{y_t}(y_{t-1})-\alpha^{\frac{|y_t|+y_t}{2}}\beta^{\frac{|y_t|-y_t}{2}}(1-\alpha-\beta)^{1-|y_t|}] = 0,
\end{aligned}$$

where $A = p\delta_{y_t}(y_{t-1}) + (1-p)\alpha^{\frac{|y_t|+y_t}{2}}\beta^{\frac{|y_t|-y_t}{2}}(1-\alpha-\beta)^{1-|y_t|}$.

The likelihood function is

$$\begin{aligned}
L &= \prod_{t=2}^{n} f(y_t|y_{t-1})f(y_1) \\
&= \alpha^{\frac{|y_t|+y_t}{2}}\beta^{\frac{|y_t|-y_t}{2}}(1-\alpha-\beta)^{1-|y_t|}\prod_{t=2}^{n}[p\delta_{y_t}(y_{t-1}) + (1-p)\alpha^{\frac{|y_t|+y_t}{2}}\beta^{\frac{|y_t|-y_t}{2}}(1-\alpha-\beta)^{1-|y_t|}].
\end{aligned}$$

The maximum likelihood estimators $\hat{\alpha}_{ML}$, $\hat{\beta}_{ML}$ and $\hat{p}_{ML}$ are obtained by solving the following nonlinear equations

$$\begin{aligned}
\frac{\partial logL}{\partial\alpha} &= \frac{|y_1|+y_1}{2\alpha} - \frac{1-|y_1|}{1-\alpha-\beta} \\
&\quad + \sum_{t=2}^{n}\frac{1}{A}(1-p)\alpha^{\frac{|y_t|+y_t}{2}-1}\beta^{\frac{|y_t|-y_t}{2}}(1-\alpha-\beta)^{-|y_t|}[\frac{|y_t|+y_t}{2}(1-\beta)+(\frac{|y_t|-y_t}{2}-1)\alpha] = 0, \\
\frac{\partial logL}{\partial\beta} &= \frac{|y_1|-y_1}{2\beta} - \frac{1-|y_1|}{1-\alpha-\beta} \\
&\quad + \sum_{t=2}^{n}\frac{1}{A}(1-p)\alpha^{\frac{|y_t|+y_t}{2}}\beta^{\frac{|y_t|-y_t}{2}-1}(1-\alpha-\beta)^{-|y_t|}[\frac{|y_t|-y_t}{2}(1-\alpha)+(\frac{|y_t|+y_t}{2}-1)\beta] = 0, \text{ and} \\
\frac{\partial logL}{\partial p} &= \sum_{t=2}^{n}\frac{1}{A}[\delta_{y_t}(y_{t-1})-\alpha^{\frac{|y_t|+y_t}{2}}\beta^{\frac{|y_t|-y_t}{2}}(1-\alpha-\beta)^{1-|y_t|}] = 0,
\end{aligned}$$

where $A = p\delta_{y_t}(y_{t-1}) + (1-p)\alpha^{\frac{|y_t|+y_t}{2}}\beta^{\frac{|y_t|-y_t}{2}}(1-\alpha-\beta)^{1-|y_t|}$.

4. Simulation Studies

To validate our results, we simulate 1000 samples of sizes $n = 100, 200$ and 500 from the Markov model of $BerD(\alpha_t, \beta_t)$ and the Jacobs-Lewis mixture model. For the Markov model, we assigned the following combinations of the parameters

$$(\delta_{00}, \delta_{01}, \delta_{10}, \delta_{11}) \in \{(1, -1, 1, 1), (-1, 2, -1, 2), (1, 2, 1, 2), (-1, -2, -1, -2)\}.$$

For the Jacobs-Lewis mixture model, the parameters selected are

$$(\alpha, \beta) \in \{(0.2, 0.2), (0.2, 0.5), (0.4, 0.1)\}, p = 0.3, 0.7.$$

We used the bias and the mean square error (MSE) as performance measures of the estimates, Tables B1 and B2 in Appendix B illustrate some of the results. For the Markov Model we found that the MSE and bias of each parameter are reciprocally related to the sample sizes as expected (except in two cases where the bias increase out of 48 cases). Also, we noticed that in most cases the bias has similar sign as the initial parameter. For the Jacobs-Lewis Mixture Model we found that the MSE of each parameter is reciprocally related to the sample size (except in one case of Yule-Walker method out of 162 cases). The MSE of CML and ML estimates are compatible. The ML estimates have smaller MSE followed by CML and last are the Yule-Walker. As p decreses the MSE of p decreases when $\beta > \alpha$, while as p increases the MSE of p decreases when $\alpha \neq \beta$. The MSE of α and β decreases with small values of p. Regarding the bias it can be seen that, p is biased negatively except in one case. In addition, the magnitude of the biases of p using ML method is less than that of the two other methods except in two cases. Hence, in terms of bias, p is better estimated using ML method while for α and β, nothing can be said. The simulation study demonstrates the efficiency and consistency of the estimates.

5. Application (SABIC Stock)

The Saudi Stock Exchange opens at 11:00 a.m. and closes at 3:30 p.m. We consider the closing price at every minute of Saudi Basic Industry (SABIC) stock on February 22, 2012. Missing minutes have been added with a zero price change. We deleted the first and final 15 minutes of the trading day in order to study the price formation during ordinary trading. SABIC share price varies in increments of ± 0.25 SAR. Hence, the tick size is 0.25. The price change is therefore characterized by discrete jumps. The frequency distribution for the difference series is presented in Table 1. The price change on that day took the values -0.5, -0.25, 0 and 0.25. The value of -0.5 occurred only once.

Table 1. Frequency distribution for the difference of SABIC.

Difference of SABIC	Count	Percent
-0.5	1	0.42
-0.25	46	19.25
0	143	59.83
0.25	49	20.50

To fit the data using our models discussed in previous sections, we set Y_t equal -1, 0 and 1 if the price change in the t^{th} minute is less than 0, equal 0 and greater than 0, respectively.

In order to get some information about the data, Tables 2 and 3 display the frequency distribution and some descriptive statistics for the transformed variable Y_t respectively. It can be seen from Table 2 that we have almost symmetric distribution with excess of zero.

Table 2. Frequency distribution for Y_t.

Y_t	Count	Percent
-1	47	19.67
0	143	59.83
1	49	20.50

Table 3. Descriptive statistics for SABIC and Y_t.

Variable	No. of observations	Mean	Standard deviation	Median	Range
SABIC	240	100.94	0.387	101	1.75
Y_t	239	0.0126	0.6383	0	2

The time series plot of the transformed stock is exhibited in Figure 1. The autocorrelation function (ACF) for Y_t is shown in Figure 2.

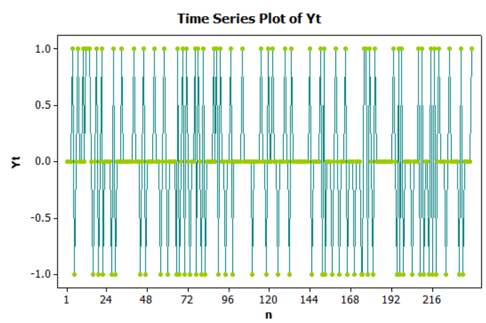

Figure 1. Time series plot for Y_t.

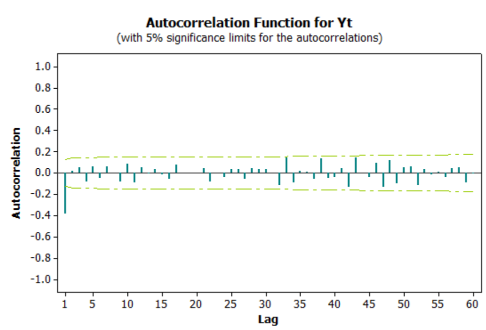

Figure 2. ACF for Y_t.

Figure 1 shows that the process is reasonably stationary in mean and variance. From Figure 2, we note that the only significant correlation is that of order one with negative correlation -0.39. Therefore, the Jacobs-Lewis Mixture model cannot be applied. The Markov model is fitted.

The estimates are obtained using maximum likelihood method of Section 2 and the standard errors are presented in Table 4. It is clear that all the parameters are significant.

Table 4. MLE and standard errors of the parameters for Markov model.

Parameter	MLE	Standard Error
δ_{00}	-1.17286	0.181604
δ_{01}	-0.76720	0.290276
δ_{10}	-1.43280	0.221337
δ_{11}	1.48363	0.324114

Thus, the estimated conditional distribution of $Y_t|Y_{t-1}$ stock is $BerD(\alpha_t, \beta_t)$, where

$$\alpha_t = \frac{e^{-1.17-0.77y_{t-1}}}{1+e^{-1.17-0.77y_{t-1}}+e^{-1.43+1.48y_{t-1}}}, \text{ and}$$
$$\beta_t = \frac{e^{-1.43+1.48y_{t-1}}}{1+e^{-1.17-0.77y_{t-1}}+e^{-1.43+1.48y_{t-1}}}.$$

Therefore, we can write the estimated probability transition matrix as

$$\begin{bmatrix} 0.032 & 0.580 & 0.389 \\ 0.154 & 0.645 & 0.200 \\ 0.479 & 0.456 & 0.065 \end{bmatrix}$$

The transition matrix reveals that the probability of visiting zero in the next transition is the highest (almost greater than 0.5). On the other hand, the transition matrix also indicates the inverse relation between two successive times as the off diagonal probabilities are higher. Using this transition matrix, we get the predicted values of the change in SABIC prices as

$$E(Y_t|Y_{t-1}=-1)=0.356,\ E(Y_t|Y_{t-1}=0)=0.047 \text{ and } E(Y_t|Y_{t-1}=1)=-0.414$$

Taking the limit of the power of this transition matrix, we get the corresponding estimated limiting distribution as $P(Y=-1)=0.198$, $P(Y=0)=0.593$ and $P(Y=1)=0.209$. This means on the long run the probability that the SABIC price will not change is almost 0.6. The distribution is symmetric with the probability of increase or decrease price change equal 0.2. From Table 2 we can see that the empirical distribution and the limiting distribution of Y_t are almost identical.

6. Conclusions

In this paper we developed two time series models based on Bernoulli Difference distribution. The Bernoulli Difference Markov model allows positive and negative correlation, while the Jacob-Lewis mixture model allows positive correlation only. The methods of estimation discussed are compatible. Many real life applications can be fitted to these models such as price change in stock market.

Acknowledgements

This project was funded by the National Plan for Science, Technology and Innovation (MAARIFAH), King Abdulaziz City for Science and Technology, Kingdom of Saudi Arabia, award number (11-MAT1856-02).

Appendix A.

Proof of Proposition 3.1

(1)
$$\begin{aligned}
E(Y_t) &= E(U_t)E(Y_{t-1}) + E(1-U_t)E(\varepsilon_t)\\
&= pE(Y_{t-1}) + (1-p)(\alpha-\beta)\\
&= p^2E(Y_{t-2}) + p(1-p)(\alpha-\beta) + (1-p)(\alpha-\beta)\\
&\vdots\\
&= p^t(\alpha-\beta) + (1-p)(\alpha-\beta)\sum_{k=0}^{n-1} p^k, \quad 0<p<1\\
&= p^t(\alpha-\beta) + (1-p)(\alpha-\beta)\frac{1-p^t}{1-p}\\
&= \alpha-\beta.
\end{aligned}$$

(2) $V(Y_t) = E(Y_t^2) - E^2(Y_t)$. The second moments is

$$\begin{aligned}
E(Y_t^2) &= E(U_t^2)E(Y_{t-1}^2) + E(1-U_t)^2E(\varepsilon_t^2) + 2E(Y_{t-1})E(U_t - U_t^2)E(\varepsilon_t)\\
&= pE(Y_{t-1}^2) + (1-p)(\alpha+\beta)\\
&= p^2E(Y_{t-2}^2) + p(1-p)(\alpha+\beta) + (1-p)(\alpha+\beta)\\
&\vdots\\
&= p^t(\alpha+\beta) + (1-p)(\alpha+\beta)\sum_{k=0}^{t-1} p^k \quad ,0<p<1\\
&= p^t(\alpha+\beta) + (1-p)(\alpha+\beta)\frac{1-p^t}{1-p}\\
&= \alpha+\beta.
\end{aligned}$$

Therefore, $Var(Y_t) = \alpha+\beta-(\alpha-\beta)^2$.

(3) The covariance between two successive observations is given by

$$\begin{aligned}
Cov(Y_t, Y_{t-1}) &= E(Cov(Y_t, Y_{t-1}|Y_{t-1})) + Cov(E(Y_t|Y_{t-1}), E(Y_{t-1}|Y_{t-1})),\\
&= Cov(E(Y_t|Y_{t-1}), Y_{t-1}) = Cov(pY_{t-1} + (1-p)(\alpha-\beta), Y_{t-1}),\\
&= pVar(Y_{t-1}) = p(\alpha+\beta-(\alpha-\beta)^2).
\end{aligned}$$

Appendix B.

Table B1. Bias and MSE for Bernoulli difference Markov model for different assumed parameters.

n	Bias(δ_{00})	Bias(δ_{01})	Bias(δ_{10})	Bias(δ_{11})	MSE(δ_{00})	MSE(δ_{01})	MSE(δ_{10})	MSE(δ_{11})
100	0.1303	-0.0152	0.0905	0.0411	0.2623	0.2742	0.2589	0.2800
200	0.0498	-0.0185	0.0300	0.0098	0.0846	0.0907	0.0832	0.0911
500	0.0135	-0.0070	0.0037	0.0020	0.0301	0.0334	0.0287	0.0332

$(\delta_{00}, \delta_{01}, \delta_{10}, \delta_{11}) = (1, -1, 1, 1)$

n	Bias(δ_{00})	Bias(δ_{01})	Bias(δ_{10})	Bias(δ_{11})	MSE(δ_{00})	MSE(δ_{01})	MSE(δ_{10})	MSE(δ_{11})
100	-0.0338	0.0374	-0.0712	0.1427	0.0958	1.0859	0.1184	1.1579
200	-0.0156	-0.0034	-0.0329	0.0307	0.0466	0.1526	0.0489	0.1604
500	-0.0046	-0.0020	-0.0116	0.0015	0.0178	0.0567	0.0175	0.0563

$(\delta_{00}, \delta_{01}, \delta_{10}, \delta_{11}) = (-1, 2, -1, 2)$

n	Bias(δ_{00})	Bias(δ_{01})	Bias(δ_{10})	Bias(δ_{11})	MSE(δ_{00})	MSE(δ_{01})	MSE(δ_{10})	MSE(δ_{11})
100	0.0842	0.1068	0.0739	0.1490	0.3798	0.4420	0.3848	0.4707
200	0.0294	0.0346	0.0227	0.0513	0.0998	0.1315	0.0975	0.1354
500	0.0211	0.0276	0.0165	0.0312	0.0352	0.0473	0.0342	0.0484

$(\delta_{00}, \delta_{01}, \delta_{10}, \delta_{11}) = (1, 2, 1, 2)$

n	Bias(δ_{00})	Bias(δ_{01})	Bias(δ_{10})	Bias(δ_{11})	MSE(δ_{00})	MSE(δ_{01})	MSE(δ_{10})	MSE(δ_{11})
100	0.0049	-0.1181	-0.0015	-0.0388	0.1040	0.7505	0.0988	0.7899
200	0.0028	-0.0482	-0.0025	-0.0226	0.0479	0.1651	0.0468	0.1735
500	-0.0003	-0.0139	0.0006	-0.0023	0.0191	0.0579	0.0170	0.0589

$(\delta_{00}, \delta_{01}, \delta_{10}, \delta_{11}) = (-1, -2, -1, -2)$

Table B2. MSE and Bias results for Jacobs-Lewis Mixture model $(\alpha, \beta, p) = (0.2, 0.5, 0.3)$.

		Bias			MSE		
Parameter	n	YW	CML	ML	YW	CML	ML
p	100	-0.01733	-0.0155	-0.01514	0.01366	0.00771	0.00769
	200	-0.00699	-0.00589	-0.00567	0.00679	0.003732	0.003734
	500	-0.00201	-0.00211	-0.00203	0.00238	0.001373	0.001373
α	100	0.00308	0.00012	0.00089	0.00299	0.002827	0.002811
	200	0.00259	0.001249	0.001725	0.00157	0.001513	0.001501
	500	0.00142	0.000955	0.001139	0.00058	0.000544	0.000542
β	100	0.00367	0.00012	0.00089	0.00285	0.002827	0.002811
	200	0.00088	0.001249	0.001725	0.00152	0.001513	0.001501
	500	0.00097	0.000955	0.001139	0.00061	0.000544	0.000542

References

Al-Osh, M. A., Alzaid A. A., 1987. First-order integer-valued autoregressive (INAR(1)) process. Journal of Time Series Analysis, 8(3), 261–275.

Alzaid, A. A., Al-Osh M. A., 1990. An integer-valued pth-order autoregressive structure (INAR(p)) process. Journal of Applied Probability, 314–324.

Alzaid, A. A., Omair, M. A., 2014. Poisson difference integer valued autoregressive model of order one. Bull. Malays. Math. Sci. Soc. (2) 37(2) (2014), 465–485.

Bulla, J., Chesneau, C., Kachour, M., 2011. A Bivariate first-order signed integer-valued autoregressive process. hal–00655102.

Jacobs, P. A., Lewis, P. A., 1978. Discrete time series generated by mixtures I: Correlational and runs properties. Journal of the Royal Statistical Society. Series B (Methodological), 94–105.

Jacobs, P. A., Lewis, P. A., 1978. Discrete time series generated by mixtures II: Asymptotic properties. Journal of the Royal Statistical Society. Series B (Methodological), 222–228.

Jacobs, P. A., Lewis, P. A., 1983. Stationary discrete autoregressive-moving average time series generated by mixtures. Journal of Time Series Analysis, 4(1), 19–36.

Kachour, M., Truquet, L., 2011. A p-order signed integer-valued autoregressive (SINAR (p)) model. Journal of Time Series Analysis, 32(3), 223–236.

Karlis, D., Anderson, J., 2009. Time series process on Z: A ZINAR model. WINTS 2009, Aveiro.

Kim, H. Y., Park, Y., 2008. A non-stationary integer-valued autoregressive model. Statistical papers, 49(3), 485–502.

McKenzie, E., 1986. Autoregressive moving-average processes with negative-binomial and geometric marginal distributions. Advances in Applied Probability, 679–705.

Odhah, O. H., 2013. Probability distributions on the class of integer numbers. Doctoral dissertation, King Saud University, 2013.

Omair, M. A., Alzaid, A., Odhah, O., 2016. A trinomial difference distribution. Revista Colombiana de Estadística, 39(1), 1–15.

Watier, L., Richardson, S., Hubert, B., 1991. A time series construction of an alert threshold with application to S. bovismorbificans in France. Statistics in Medicine, 10(10), 1493–1509.

Zhang, H., Wang, D., Zhu, F., 2010. Inference for INAR(p) processes with signed generalized power series thinning operator. Journal of Statistical Planning and Inference, 140(3), 667–683.