Journal of Statistical Modelling: Theory and Applications
Vol. 2, No. 2, 2021, pp. 51-62
Yazd University Press 2021

JSMTA

Research Paper

On the estimation problem in AR(1) model with exponential innovations

Abdollah Saadatmand*[1], Ali Reza Nematollahi[2],
Soltan Mohammad Sadooghi-Alvandi[2]
[1]Department of Statistics, College of Science, Payame Noor University,
P.O. Box, 19395-4697, Tehran, Iran
[2]Department of Statistics, College of Science, Shiraz University, Shiraz, Iran

Received: March 19, 2022/ Revised: April 30, 2022 / Accepted: May 02, 2022

Abstract: In this article, the autoregressive model of order one with exponential innovations is considered. The maximum likelihood and Bayes estimators of the autoregression parameter, under squared error loss function with non-informative prior are examined. A simulation study is conducted to compare the behavior of the estimators via their relative bias and risks. Moreover, a real data example is presented.

Keywords: Autoregressive model; Bayes estimation; Exponential innovations; Maximum Likelihood estimation.

1 Introduction

In many statistical studies especially in time series models, the observed values of the considered variables are nonnegative and for analyzing such observations it is convenient to use the models with nonnegative innovations such as exponential innovations. The several attractive features of the resulting exponential autoregressive process can be found in Gaver and Lewis (1980). They showed that $\{X_j\}$ has an exponential marginal distribution with parameter λ, denoted by Exp (λ), provided that ρ should be positive, $0 \leq \rho < 1$ and

$$X_j = \begin{cases} \rho X_{j-1} & \text{with prob. } \rho \\ \rho X_{j-1} + \varepsilon_j & \text{with prob. } 1-\rho \end{cases}$$

*Corresponding author: `a.saadatmand@pnu.ac.ir`

where ε_j ' s are i.i.d. Exp (λ) random variables with density:

$$f(x) = \frac{1}{\lambda} e^{\frac{-x}{\lambda}}, \quad \lambda > 0.$$

However, as noted by Bell and Smith (1986), this scheme is not an AR(1) process. Bell and Smith (1986) have studied an AR(1) model $X_j = \rho X_{j-1} + \varepsilon_j$ with $0 < \rho < 1$ where ε_j is a nonnegative white noise with a finite second moment and have applied this model to water quality analysis. They studied the inference problem for the Gaussian, uniform, and exponential distributions, and a nonparametric case with positive continuous distributed white noise. It was proved that the maximum likelihood estimator (MLE) has positive bias and is a strongly consistent estimator of ρ unless there exist constants c and d such that $0 < c < d < \infty$ with $F(d) - F(c) = 1$, where $F(.)$ is distribution function of ε_j.

A mean stationary AR(1) model with exponential innovation was considered by Anděl (1988), where he derived the distribution of an approximation to the MLE for the autoregressive parameter. These results were generalized to mean stationary AR(2) by Anděl (1989). Moreover, generalization to mean stationary vector AR(1) models and strong consistency of estimators were studied by Anděl (1992) and Anděl (1998). Anděl and Garrido (1991) studied the nonnegative AR(2) processes in a Bayesian setting. Datta and McCormick (1995) constructed a confidence interval for the autoregressive parameter of the AR(1) process with positive innovations. The class of continuous-time stochastic volatility models for financial assets are studied by Nielsen and Shephard (2001), Brockwell and Marquardt (2003), and Brockwell and Marquardt (2005). In these kinds of models, the volatility processes are the solutions to the Ornstein-Uhlenbeck (OU) processes driven by non-decreasing Levy processes, whose paths are nonnegative whenever the kernel is nonnegative. A necessary and sufficient condition for the kernel to be nonnegative in terms of its Laplace transform is established by Tsai and Chan (2005). Similar conditions are also derived by Tsai and Chan (2007) for the discrete-time autoregressive moving average (ARMA) processes in terms of kernel generating function.

The Bayesian estimation and prediction for the mean stationary autoregressive model AR(1) with exponential innovations under the square error loss function were studied by Ibazizen and Fellag (2003). They generalized Turkmann's (1990) results by using the more general non-informative priors and compared the results with Anděl (1988) and Turkmann (1990). Nielsen and Shephard (2003) derived the exact and asymptotic distributions of the MLE for the autoregression parameter of the AR(1) model with exponential innovations. They showed that MLE is consistent with different rates for different values of autoregression parameters. Larbi and Fellag (2016) performed a robust Bayesian analysis of the Bayesian estimation of a mean stationary autoregressive model with exponential innovations and obtained optimal Bayesian estimators of the parameters corresponding to the smallest oscillation of the posterior risks. However, all of their estimators involve Appell hypergeometric functions in their formulas and it is complicated in practice. In another study on this model, Saadatmand et al. (2017) considered the estimation of a missing value for the stationary AR(1) model with exponential innovations introduced by Anděl (1988) and compared two methods of estimation of the missing value with respect to Pitman's measure of closeness.

In this paper, we consider the first order autoregressive model with exponential error introduced by Bell and Smith (1986) as

$$X_j = \rho X_{j-1} + \varepsilon_j, \quad 0 < \rho < 1, \quad j = 1, ..., n,$$

where ε_j ' s are i.i.d. Exp (λ) random variables and X_0 is fixed. Note that this model is not mean stationary and differs from the model which has been considered by Ibazizen and Fellag (2003) and Larbi and Fellag (2016). The parameter of interest is ρ and the nuisance parameter is λ. We have compared the competing estimation procedures; MLE and Bayes estimation. In Section 2, we review the MLE introduced in Bell and Smith (1986). In Section 3, we obtain the Bayes estimator under the square error loss function. In Section 4, a simulation study is conducted to compare the behavior of the estimators via relative bias and frequentist risk, the expected loss of both the data, and unknown parameters. Finally, in Section 5, a real data example is presented.

2 Maximum likelihood estimation

We consider the first-order autoregressive model with exponential error defined by:

$$X_j = \rho X_{j-1} + \varepsilon_j, \quad 0 < \rho < 1, \quad j = 1, ..., n, \tag{1}$$

where $X_0 \geq 0$ is fixed and $\varepsilon_j^{'}$s are independent identically distributed Exp (λ) random variables such that the joint probability density function (pdf) of $(X_1, ..., X_n|\lambda, \rho)$ is given by

$$\begin{aligned} f_{\boldsymbol{X}}(x_1, \ldots, x_n|\lambda, \rho) &= f_{\boldsymbol{\varepsilon}}(x_1 - \rho x_0, x_2 - \rho x_1, ..., x_n - \rho x_{n-1}|\lambda, \rho), \\ &= \frac{1}{\lambda^n} e^{-\frac{1}{\lambda}\sum_{j=1}^{n}(x_j - \rho x_{j-1})} I_{(0, \min_{1\leq j\leq n}(1, \frac{x_j}{x_{j-1}}))}(\rho) I_{(0,\infty)}(\lambda). \end{aligned} \tag{2}$$

Note that the joint pdf is increasing in ρ, and therefore the MLE of ρ (as given by Bell and Smith, 1986, in nonparametric case) is

$$\hat{\rho}_{ml} = \min_{1\leq j\leq n}(1, \frac{x_j}{x_{j-1}}). \tag{3}$$

The consistency of $\hat{\rho}_{ml}$ can be proved in the similar manner to Bell and Smith (1986). Furthermore, the MLE of λ (similar to Bell and Smith, 1986) can be obtained from the likelihood function (2) as

$$\hat{\lambda}_{ml} = \frac{\sum_{j=1}^{n}(x_j - \hat{\rho}_{ml} x_{j-1})}{n}. \tag{4}$$

Since $e_j = X_j - \hat{\rho}_{ml} X_{j-1}$ is the estimate of ε_j, from the consistency of $\hat{\rho}_{ml}$ and applying strong law of large numbers, it can be shown that $\hat{\lambda}_{ml}$ as the mean of e_j ' s, is a consistent estimator for λ, which is the expected value of $\varepsilon_j^{'}$s.

3 Bayes estimation

To apply a Bayes estimation procedure, we consider improper (non-informative) prior for (ρ, λ) as

$$\pi(\rho, \lambda) \propto \frac{const.}{\lambda}, \quad 0 < \rho < 1, \quad \lambda > 0. \tag{5}$$

Our reasons for introducing this improper prior are simplicity, having closed-form and optimality of the resulted Bayes estimator in comparison to the other estimators obtained from some of considered competitor priors. Our investigation of other priors is huge and not reported in this paper. Then from (2) and (5), we have

$$f_{\boldsymbol{X}}(x_1, \ldots, x_n; \lambda, \rho) = const. \frac{1}{\lambda^{n+1}} e^{-\frac{1}{\lambda}\sum_{j=1}^{n}(x_j - \rho x_{j-1})} I_{(0,\hat{\rho}_{ml}))}(\rho) I_{(0,\infty)}(\lambda).$$

It is clear that

$$f(\rho, \lambda | x_1, \ldots, x_n) \propto \frac{1}{\lambda^{n+1}} e^{-\frac{1}{\lambda}\sum_{j=1}^{n}(x_j - \rho x_{j-1})} I_{(0,\hat{\rho}_{ml}))}(\rho) I_{(0,\infty)}(\lambda).$$

Let $A = \sum_{j=1}^{n} x_j$, $B = \sum_{j=1}^{n} x_{j-1}$ and $C = A - B\hat{\rho}_{ml}$. Then it can be shown that the posterior pdf of (ρ, λ) given $\boldsymbol{x} = (x_1, \ldots, x_n)$ is

$$f(\rho, \lambda | \boldsymbol{x}) = \frac{BC^{n-1}}{(n-2)![1 - (\frac{C}{A})^{n-1}]} \frac{1}{\lambda^{n+1}} e^{-\frac{1}{\lambda}(A - B\rho)} I_{(0,\hat{\rho}_{ml}))}(\rho) I_{(0,\infty)}(\lambda),$$

and consequently,

$$f(\rho | \boldsymbol{x}) = \frac{BC^{n-1}}{[1 - (\frac{C}{A})^{n-1}]} \frac{(n-1)}{(A - B\rho)^n} I_{(0,\hat{\rho}_{ml}))}(\rho).$$

The Bayes estimator under the square error loss function $l(\hat{\rho}, \rho) = (\hat{\rho} - \rho)^2$ is posterior mean. Then

$$\begin{aligned} \hat{\rho}_{bs} &= E(\rho|\boldsymbol{x}) = \int_0^{\hat{\rho}_{ml}} \rho . f(\rho|\boldsymbol{x}) d\rho = \int_0^{\hat{\rho}_{ml}} \rho . \frac{BC^{n-1}}{[1 - (\frac{C}{A})^{n-1}]} \frac{(n-1)}{(A - B\rho)^n} d\rho \\ &= \frac{\hat{\rho}_{ml} - \frac{C}{B(n-2)}[1 - (\frac{C}{A})^{n-2}]}{1 - (\frac{C}{A})^{n-1}} \end{aligned} \tag{6}$$

Proposition 3.1. *The estimator $\hat{\rho}_{bs}$ in* (6) *is a consistent estimator of ρ.*

Proof. It is obvious that $\frac{A}{B} \to 1$ as $n \to \infty$, such that

$$\frac{C}{A} = 1 - \frac{B}{A}\hat{\rho}_{ml} \to 1 - \rho, \qquad \frac{C}{B} = \frac{A}{B} - \hat{\rho}_{ml} \to 1 - \rho.$$

The consistency of $\hat{\rho}_{bs}$ obtained by the consistency of $\hat{\rho}_{ml}$ as

$$\hat{\rho}_{bs} = \frac{\hat{\rho}_{ml} - \frac{C}{B(n-2)}[1 - (\frac{C}{A})^{n-2}]}{1 - (\frac{C}{A})^{n-1}} \to \frac{\rho - \frac{1-\rho}{\infty}[1 - 0]}{1 - 0} = \rho.$$

□

Similarly, the Bays estimator λ can be obtained as $\hat{\lambda}_{bs} = \frac{C}{(n-2)} \frac{1-(\frac{C}{A})^{n-2}}{1-(\frac{C}{A})^{n-1}}$. But this estimator has no advantage to $\hat{\lambda}_{ml}$ in (2.4) and we do not apply $\hat{\lambda}_{bs}$ for our predictions in Section 5. In the next two Sections, we compare $\hat{\rho}_{ml}$ and $\hat{\rho}_{bs}$ by simulation studies and in a real data example.

4 Simulation studies

In the following, for each value of sample size n, we simulate innovations as a sample from exponential distribution with expectation λ to generate 100000 replications of AR(1) model with initial value X_0 and autoregression parameter ρ . We study the values $\lambda = 0.1, 0.2, 0.5, 1, 2, 5, 10, 20, 50, 100$, different values $X_0 = 0, 0.1, 0.2, 0.5, 1, 2, 5, 10, 20, 50, 100$ for sample sizes $n = 5, 7, 10, 15, 20, 30, 50$ and $0 \leq \rho \leq 1$ (by steps 0.01) and report a subset of these values. For each replication, we obtain the MLE (3) and the Bayes estimator (6) and calculate their square error loss function $(\hat{\rho} - \rho)^2$ and their relative bias $\frac{(\hat{\rho}-\rho)}{\rho}$.

Figures 1 to 5 show the risk of the MLE and Bayes estimator, for some values of n, λ and X_0, as the mean of their square error loss functions over all replications. Similarly, Figures 6 to 10 show the relative bias of MLE and Bayes estimators for different values of n, λ and X_0, as the average of their relative bias over all replications.

According to Figures 1 to 10, the risk and relative bias decrease as the sampler size n increases (as be expected), and also for larger values of ρ, these quantities have smaller values. Furthermore, for each fixed λ, the larger X_0 leads to the smaller risk and the smaller relative bias. On the other hand, for each fixed X_0, the larger λ results in the larger risk and also the larger relative bias. Finally, the Bayes estimator has a smaller risk than the MLE for each n, λ and X_0. Moreover, it can be observed that the Bayes estimates, with smaller relative bias, have better performances than the MLE for different values of the parameters.

5 A real data example

We consider Viscosity data (series D, Box and Jenkins, 1976) which consists of 310 observations to investigate usefulness of the Bayes estimator (6). Datta and McCormick (1995) considered this data set to fit a positive AR(1) model, see Datta and McCormick (1995) for more details. We assume the first 301 observations of the series as observed values $X_0, X_1, \ldots, X_{300}$ and by using AR(1) model (1) try to predict the last 9 observations $X_{301}, X_{302}, \ldots, X_{309}$ and compare the resulted values with the observed values. For predicting the next steps, under the square loss function, we obtain the best predictor of X_k based on the previous values $X_0, X_1, \ldots, X_{k-1}$, i. e., $\hat{X}_k = E(X_k|X_0, X_1, \ldots, X_{k-1})$. See Saadatmand et al. (2017) for more details and discussions. Since ρ and λ are unknown, we use equation (8) in Saadatmand et al. (2017) to obtain the predictions for $k = 301, \ldots, 309$ as

$$\hat{X}_k(ml) = \hat{\rho}_{ml}\hat{X}_{k-1}(ml) + \hat{\lambda}_{ml}, \tag{7}$$

$$\hat{X}_k(bs) = \hat{\rho}_{bs}\hat{X}_{k-1}(bs) + \hat{\lambda}_{ml}, \tag{8}$$

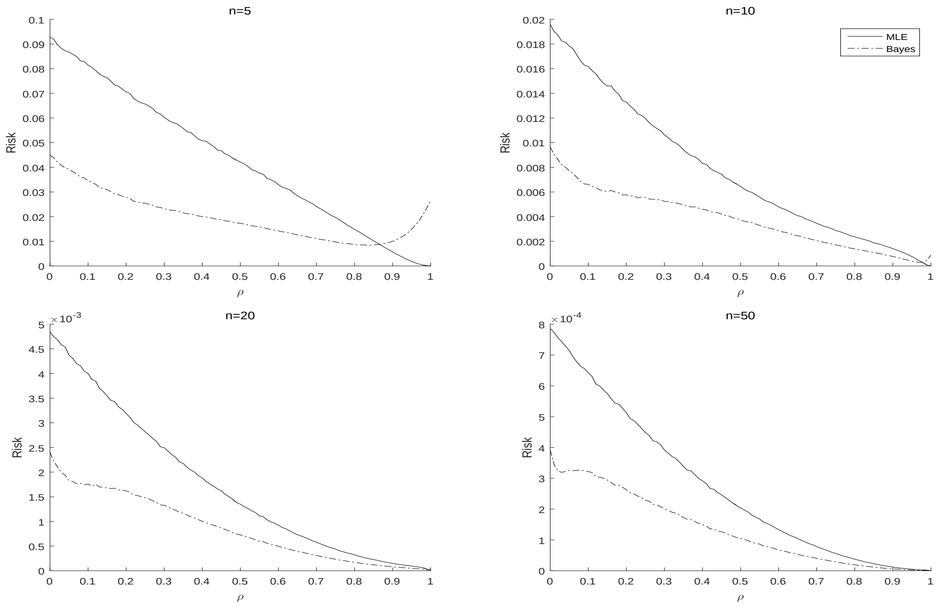

Figure 1: Risk of the MLE and Bayes estimators $X_0 = 0, \lambda = 1$ and $0 \leq \rho \leq 1$.

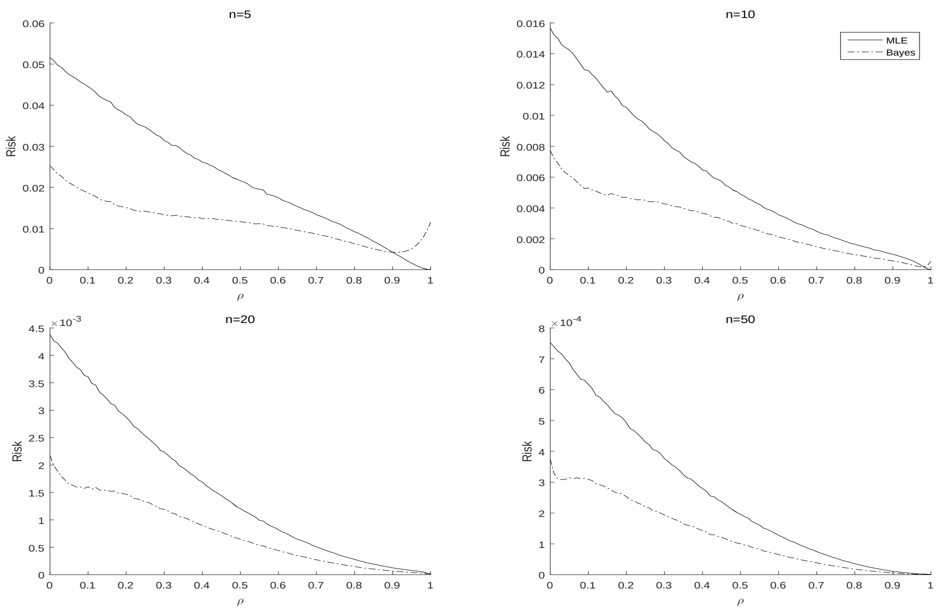

Figure 2: Risk of the MLE and Bayes estimators $X_0 = 1, \lambda = 1$ and $0 \leq \rho \leq 1$.

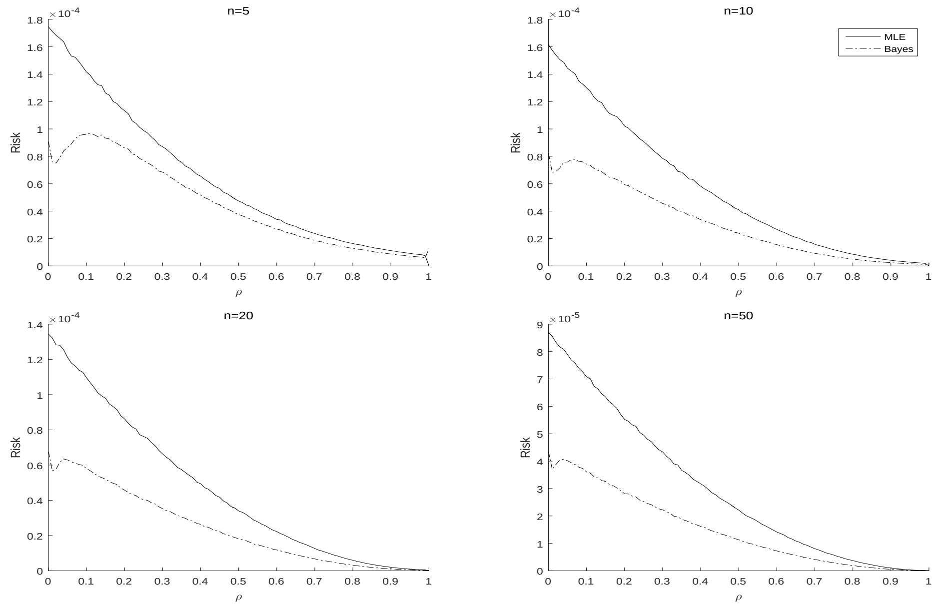

Figure 3: Risk of the MLE and Bayes estimators $X_0 = 100, \lambda = 1$ and $0 \leq \rho \leq 1$.

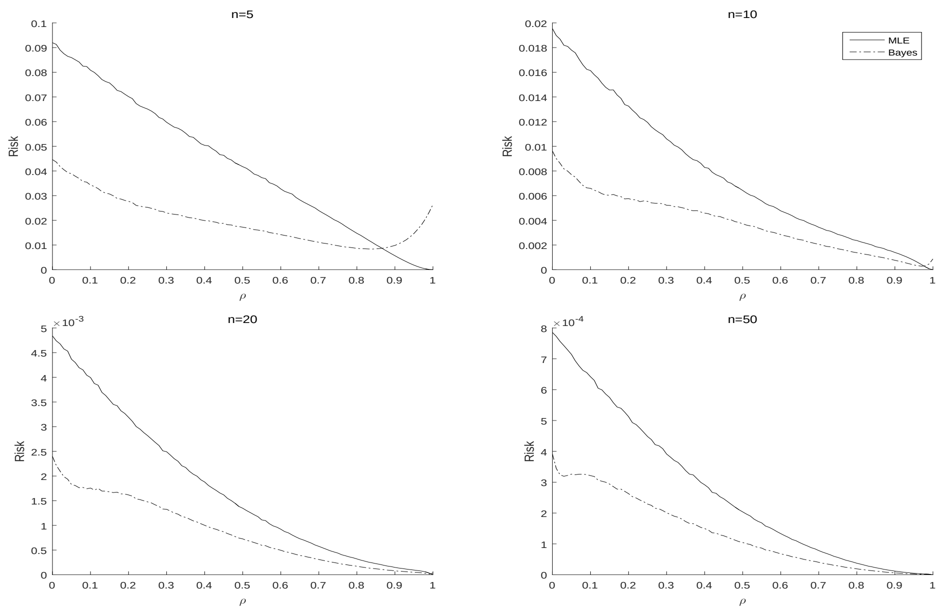

Figure 4: Risk of the MLE and Bayes estimators $X_0 = 1, \lambda = 100$ and $0 \leq \rho \leq 1$.

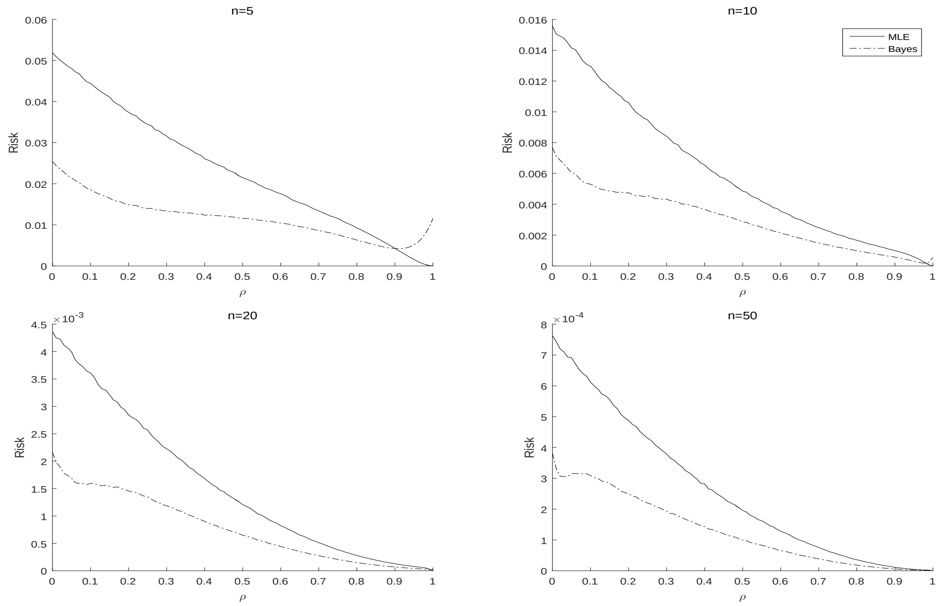

Figure 5: Risk of the MLE and Bayes estimators $X_0 = 100, \lambda = 100$ and $0 \leq \rho \leq 1$.

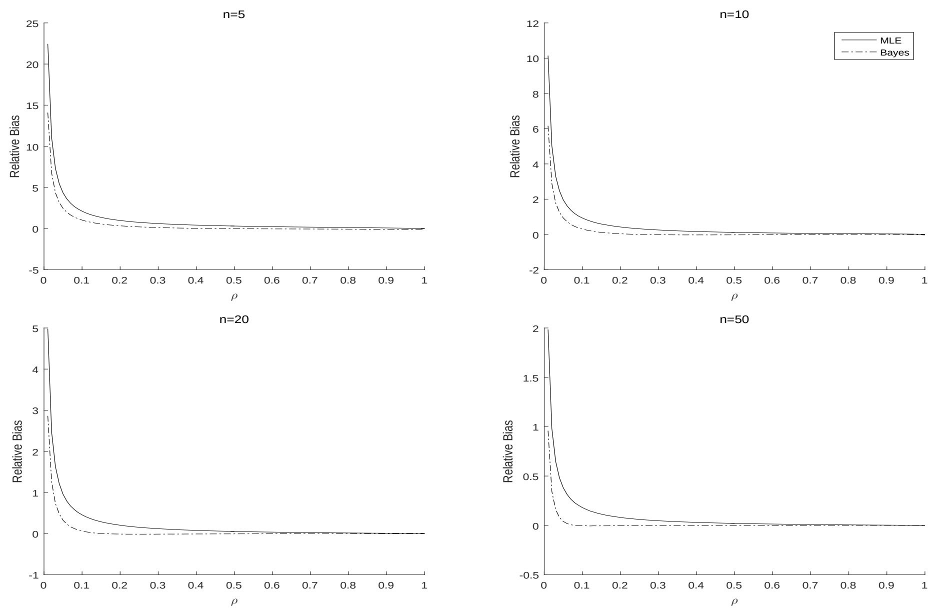

Figure 6: Relative bias of the MLE and Bayes estimators $X_0 = 0, \lambda = 1$ and $0 \leq \rho \leq 1$.

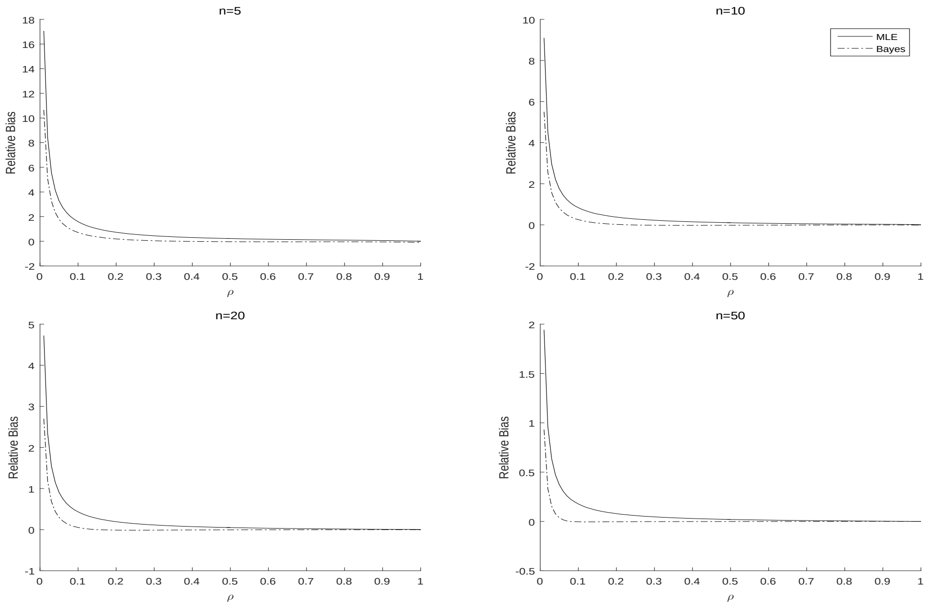

Figure 7: Relative bias of the MLE and Bayes estimators $X_0 = 1, \lambda = 1$ and $0 \leq \rho \leq 1$.

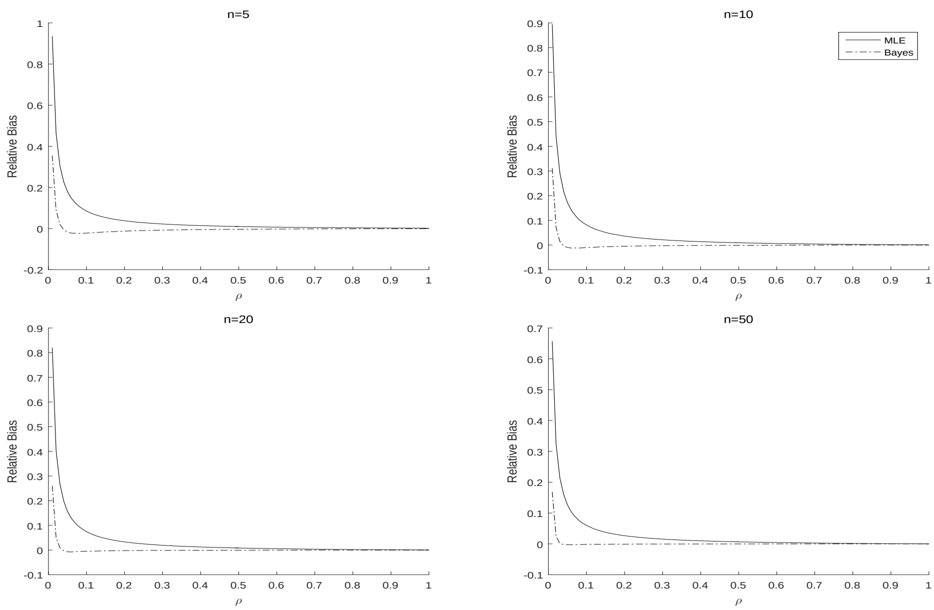

Figure 8: Relative bias of the MLE and Bayes estimators $X_0 = 100, \lambda = 1$ and $0 \leq \rho \leq 1$.

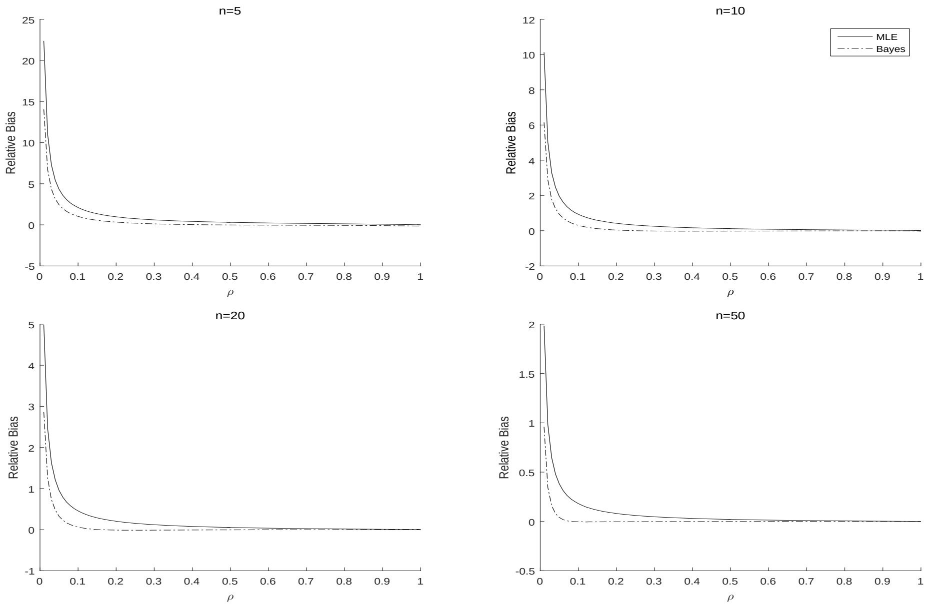

Figure 9: Relative bias of the MLE and Bayes estimators $X_0 = 1, \lambda = 100$ and $0 \leq \rho \leq 1$.

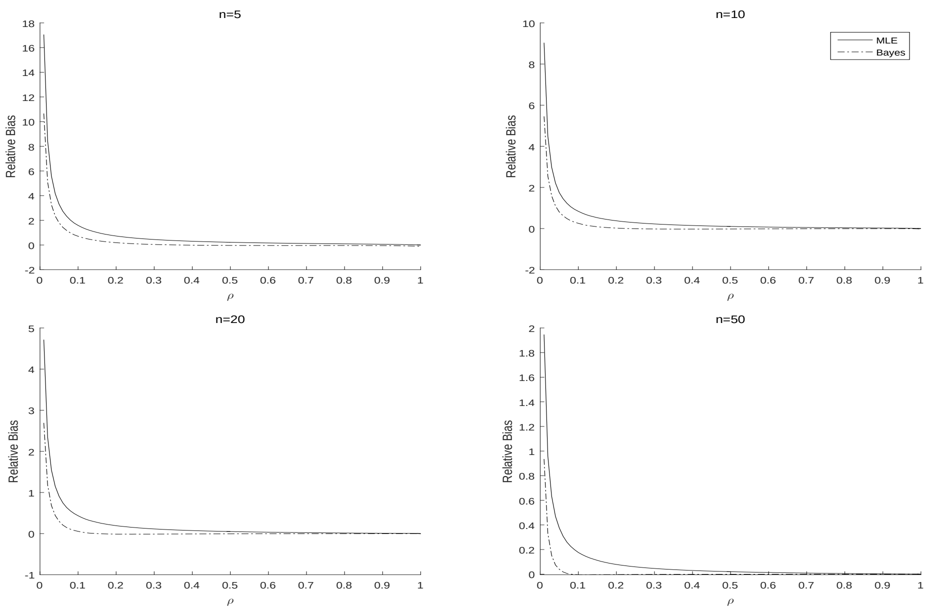

Figure 10: Relative bias of the MLE and Bayes estimators $X_0 = 100, \lambda = 100$ and $0 \leq \rho \leq 1$.

where $\hat{X}_{300}(ml) = \hat{X}_{300}(bs) = X_{300}$ and $\hat{\rho}_{ml}$, $\hat{\lambda}_{ml}$ and $\hat{\rho}_{bs}$ are given in (3), (4) and (6) and turned out to be 0.86, 1.283 and 0.8595, respectively. Table 1 presents the last 9 observed values $X_{301}, X_{302}, \ldots, X_{309}$ and their predicted values using (7) and (8).

Table 1: The observed values in Viscosity data and their predictions applying $\hat{\rho}_{ml}$ and $\hat{\rho}_{bs}$.

k	301	302	303	304	305	306	307	308	309
X_k	8.6	8.3	7.9	8.5	8.7	8.9	9.1	9.1	9.1
$\hat{X}_k(MLE)$	8.7650	8.8208	8.8689	8.9102	8.9457	8.9763	9.0026	9.0252	9.0446
$\hat{X}_k(Bayes)$	8.7609	8.8132	8.8581	8.8968	8.93	8.9586	8.9831	9.0042	9.0223

We compared these two sets of prediction values by using the criteria

$$SSE = \sum_{k=301}^{309} (\hat{X}_k - X_k)^2,$$

which can be calculated for both inference procedure so that $SSE(MLE) = 1.4898$ and $SSE(Bayes) = 1.4499$. As a result, the Bayes estimator $\hat{\rho}_{bs}$ in (6) has smaller SSE in comparison with the MLE $\hat{\rho}_{ml}$ in (3), and results in more convenient predictions for $X_{301}, X_{302}, \ldots, X_{309}$.

Acknowledgement

The authors are thankful to the two referees for their useful comments which helped to improve this paper.

References

Anděl, J. (1988). On AR(1) process with exponential white noise. *Communications in Statistics-Theory and Methods*, **17**(5):1481–1495.

Anděl, J. (1989). Non-negative autoregressive processes. *Journal of Time Series Analysis*, **10**(1):1–11.

Anděl, J. (1992). Nonnegative multivariate AR(1) processes. *Kybernetika.*, **28**(3):213–226.

Anděl, J. (1998). An estimator for parameters of a nonlinear nonnegative multidimensional AR(1) Process. *Applications of Mathematics*, **43**(5):389–398.

Anděl, J. and Garrido, M. (1991). Bayesian analysis of non-negative AR(2) processes. *Statistics*, **22**(4):579–588.

Bell, C.B. and Smith, E.P. (1986). Inference for non-negative autoregressive schemes. *Communications in Statistics- Theory and Methods*, **15**(8):2267–2293.

Box, G.E.P. and Jenkins, G.M. (1976). *Time Series Analysis: Forecasting and Control.* San Francisco: Holden-Day.

Brockwell, P.J. and Marquardt, T. (2003). *Fractional Integration of Continuous-time ARMA Processes.* Technical report of Technische Universitat Munchen.

Brockwell, P.J. and Marquardt, T. (2005). Levy-driven and fractionally integrated ARMA processes with continuous time parameter. *Statistica Sinica*, **15**:477–494.

Datta, S. and McCormick, W.P. (1995). Bootstrap inference for a first-order autoregression with positive innovations. *Journal of the American Statistical Association, Theory and Methods*, **90**(432):1289–1300.

Gaver, D.P. and Lewis, P.A.W. (1980). First-order autoregressive gamma sequences and point processes. *Advanced in Applied Probability*, **12**:27–745.

Ibazizen, M. and Fellag, H. (2003). Bayesian estimation of an AR(1) process with exponential white noise. *Statistics*, **37**(5):365–372.

Larbi, L. and Fellag, H. (2016). Robust Bayesian analysis of an autoregressive model with exponential innovations. *Afrika Statistika*, **11**(1):955–964.

Nielsen, B. and Shephard, N. (2001). Non-Gaussian Ornstein-Uhlenbeck based models and some of their uses in financial economics (with discussion). *Journal of the Royal Statistical Society, Series B*, **63**(2):167–241.

Nielsen, B. and Shephard, N. (2003). Likelihood analysis of a first-order autoregressive model with exponential innovations. *Journal of Time Series Analysis*, **24**(3):337–344.

Saadatmand, A., Nematollahi, A.R. and Sadooghi-Alvandi, S.M. (2017). On the estimation of missing values in AR(1) model with exponential innovations. *Communications in Statistics- Theory and Methods*, **46**(7):3393–3400.

Tsai, H. and Chan, K.S. (2005). A note on non-negative continuous time processes. *Journal of the Royal Statistical Society, Series B*, **67**(4):589–597.

Tsai, H. and Chan, K.S. (2007). A note on non-negative ARMA processes. *Journal of Time Series Analysis*, **28**(3):350–360.

Turkmann, M.A.A. (1990). Bayesian analysis of an autoregressive process with exponential white noise. *Statistics*, **21**(4):601–608.