Economic Computation and Economic Cybernetics Studies and Research, Issue 4/2016, Vol. 50

LecturerYakup ARI,PhD
Department of Mathematics, Yeditepe University, Istanbul, Turkey
E-mail: yakupari@gmail.com
ProfessorEmeritusAlexandros PAPADOPOULOS, PhD
Department of Mathematics,Yeditepe University, Istanbul, Turkey
E-mail: routsiag@yahoo.com

BAYESIAN ESTIMATION OF THE PARAMETERS OF THE ARCH MODEL WITH NORMAL INNOVATIONS USING LINDLEY'S APPROXIMATION

Abstract*. Autoregressive conditionally heteroscedastic (ARCH) models are used to analyze empirical financial data and capture various stylized facts in financial econometrics. The procedure that is most commonly used for estimating the unknown parameters of an ARCH model is the maximum likelihood estimation (MLE). In this study, it is assumed that the parameters of the ARCH model are random variables having known prior probability density functions, and therefore they will be estimated using Bayesian methods. The Bayesian estimators are not in a closed form, and thus Lindley's approximation will be used to estimate them. The Bayesian estimators are derived under squared error loss (SEL) and linear exponential (LINEX) loss functions. An example is given in order to illustrate the findings and furthermore, Monte Carlo simulations are performed in order to compare the ML estimates to the Bayesian ones. Finally, conclusions on the findings are given.*

Keywords: *ARCH, QML method, Lindley's Approximation, Bayesian Methods, SEL, LINEX.*

JEL Classification:62M10, 62F15, 91B84

1. Introduction

The least squares models are important in applied econometrics, because they help applied econometricians to determine how much one variable will change in response to a change in another variable. Once a model has been fitted to the data one wants to forecast and analyze the size of the errors involved in the fitting. The main assumption in the basic version of the least squares model is that the expected value of all squared error terms are the same at any given point. This assumption is called homoskedasticity. Often, the variances of the error terms are not equal and the error terms are larger for some points or ranges of the data than for others. In this situation where the variance of the error terms is not constant it is said that heteroskedasticity exists. The economist Engle (1982) treated heteroskedasticity as a variable to be modeled, and thus he constructed the first volatility model, namely the ARCH model. The basic idea of the ARCH model is that the error term a_t of an asset return is serially uncorrelated, though dependent on its p squared lag values. The dependence of a_t can be described by a simple quadratic function of its lagged values. The ARCH(p) model assumes that

$$a_t = \sigma_t \varepsilon_t\ ,\ \ \sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \alpha_2 a_{t-2}^2 + \dots. + \alpha_p a_{t-p}^2 \qquad (1)$$

where $\{\varepsilon_t\}$ is a sequence of independent and identically distributed (i.i.d.) random variables with mean zero and variance 1.

Engle (1982) used the maximum likelihood method to estimate the unknown parameters $\alpha_0, \alpha_1, \dots, \alpha_p$. Another commonly used estimation procedure for an ARCH model is the quasi maximum likelihood estimation (QMLE), whose asymptotic properties have been extensively studied, see for instance Berkes et al. (2003), Francq and Zakoïan(2004), and Straumann(2005). It should be noted that the ML and QML estimation methods require the use of a numerically optimization procedure. Bose and Mukherjee (2003) recommend a two–stage least squares estimator (LSE) for the ARCH model. These estimators are in a closed form and have the same asymptotic efficiency as those of the QMLE. Giraitis and Robinson [6] propose an estimation method that is based on Autoregressive and Autoregressive Moving Average representations of the squared ARCH process.

The ML method is the favored approach for making inferences for ARCH models. It is appealing because it is easy to implement and is available in economic packages; furthermore, Bollereslev et al. (1994), Lee and Hansen (1994) have shown that the estimators are asymptotically optimum. Another approach of making inferences about the ARCH models is the Bayesian one. However, for most of the prior distributions assumed for the model parameters, the posterior distribution is analytically intractable, and thus numerical methods or proper approximation is required. Markov chain Monte Carlo (MCMC) methods enable to draw samples from the posterior and predictive distributions, thus sample averages can be used to approximate expectations. Mitsui and Watanabe (2003) introduced a Taylored approach based on the acceptance-rejection Metropolis-Hastings algorithm for parametric ARCH-type models.

The purpose of this study is to derive Bayesian estimators for the parameters of the ARCH(p) model, using different loss functions and when the innovations are distributed according to the standard normal or a standardized student-*t* distribution. Although there do not exist any certain rules that one could follow to choose the loss function; nevertheless, this choice is of fundamental importance in Bayesian decision making. The symmetric SEL function is the usual choice due to its mathematical tractability. As it was pointed out by Moore and Papadopoulos (2000), such a choice is arbitrary and its popularity is due to its analytical tractability. In particular, Box and Tiao(1992) state that "the quadratic loss function leading to the posterior mean is arbitrary. The question remains as to why many sampling theorists seem to cling rather tenaciously to the mean squared error criterion and the quadratic loss function. As an example, banks and in general financial institutions are likely to have asymmetric preferences, as pointed out by Peel and Nobay(1998), perhaps tending to error in the direction of caution in reaching inflation targets. Consequently, in economics and finance forecasting performance is increasingly evaluated under the asymmetric loss functions as witnessed by such scholars as Batchelor and Peel (1998), Christoffersen and Diebold (1997), Granger and Newbold(1986), Granger and Pesaran(2000), West, Edison and Choi (1993)and Zellner(1986). Thus, in addition to the SEL function the asymmetric LINEX function that was introduced by Varian (1974)will be utilized.

This paper is organized as follows. In section 2, under the assumption that the innovations follow the standard normal distribution and that the parameters behave as random variables with α_0 having a gamma or vague prior and $\alpha_1, \ldots, \alpha_p$ a Dirichlet prior, the posterior density is derived which is not a closed form. Therefore, for the estimation of the parameters of the ARCH(1) and ARCH(2) models Lindley's (1980)approximation is used.. Finally, in Section 3 an example is given to illustrate the findings of section 2. In section 4, in order to compare the different types of estimators a computer simulation study is done and the results are discussed.

1. Bayesian estimation of the parameters of ARCH(p) model with normal innovations

Let $\{a_t\}$ where $t = 1,2, \ldots,$ n, denote the ARCH(p) process defined by equation (1) where the parameters $\alpha_0 > 0$ and $\alpha_i \geq 0$ for $i = 1,2, \ldots, p$. In this study it will be assumed that the process is stationary and thus the coefficients $\alpha_1, \ldots, \alpha_p$ satisfy the condition $\sum_{i=1}^{p} \alpha_i < 1$.

Let $\underline{\alpha} = (\alpha_0, \alpha_1, \ldots, \alpha_p)$ denote the parameters of the ARCH(p) model and $\underline{x} = (a_{p+1}, a_2, \ldots, a_T)$ the observed series. Then under the normality assumption, the conditional likelihood function of an ARCH(p) model is

$$f_1\left(a_{p+1}, a_{p+2}, \dots, a_T \middle| \underline{\alpha}, a_1, a_2, \dots, a_p\right) = \prod_{t=p+1}^{T} \frac{1}{\sqrt{2\pi\sigma_t^2}} e^{-\frac{a_t^2}{2\sigma_t^2}} \tag{2}$$

whereσ_t^2 can be evaluated recursively from eq (1).

The estimates obtained by maximizing eq (2) are known as the conditional maximum likelihood estimates. Usually, it is easier to maximize the log of the likelihood function, i.e.

$$L = \ln\left(\underline{x}\middle|\underline{\alpha}\right) = \sum_{t=p+1}^{T} \{-\frac{1}{2}\ln(2\pi) - \frac{1}{2} ln\left(\alpha_0 + \alpha_1 a_{t-1}^2 + \dots + \alpha_p a_{t-p}^2\right) - \frac{1}{2}\frac{a_t^2}{(\alpha_0 + \alpha_1 a_{t-1}^2 + \dots + \alpha_p a_{t-p}^2)}\} \tag{3}$$

Engle(1982) derived the MLE estimates of $\underline{\alpha}$, denoted by $\underline{\hat{\alpha}}$, by maximizing eq(3).

In this study it will be assumed that the parameters of the ARCH(p) model behave as random variables and thus they will be estimated using Bayes theorem. It will be assumed that α_0 has gamma prior $g_1(\alpha_0, r, \beta)$ with parameters (r, β),

$$g_1(\alpha_0; r, \beta) = \frac{1}{\Gamma(r)\beta^r} \alpha_0^{r-1} e^{-\frac{\alpha_0}{\beta}} \quad \text{for } \alpha_0 > 0 \text{ and } r, \beta > 0 \tag{4}$$

Furthermore, we will assume that the joint density function of $\alpha_1, \alpha_2 \dots, \alpha_p$ with $p \geq 2$ is the Dirichlet probability function with parameters $\omega_1, \omega_2, \dots, \omega_{p+1}$>0 given as

$$g_2\left(\alpha_1, \alpha_2, \dots, \alpha_p; \omega_1, \omega_2, \dots, \omega_{p+1}\right) = \frac{1}{B\left(\underline{\omega}\right)} \prod_{i=1}^{p+1} \alpha_i^{\omega_i - 1} \tag{5}$$

where$\alpha_1 + \alpha_2 + \dots + \alpha_p < 1$ and $\alpha_{p+1} = 1 - \alpha_1 + \alpha_2 + \dots + \alpha_p$. The normalizing constant $B\left(\underline{\omega}\right)$is the multinomial beta function given as

$$B\left(\underline{\omega}\right) = \frac{\prod_{i=1}^{p+1} \Gamma(\omega_i)}{\Gamma(\sum_{i=1}^{p+1} \omega_i)}$$

where $\underline{\omega} = (\omega_1 + \omega_2 + \dots + \omega_{p+1})$. If $p = 1$ the Dirichlet reduces to the beta probability function.

Since α_0 and $\alpha_1, \alpha_2, \dots, \alpha_p$ are independent, their joint pdf is given by

$$g(\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_p) = \frac{1}{\Gamma(r)\beta^r} \frac{1}{B(\underline{\omega})} \alpha_0^{r-1} e^{-\frac{\alpha_0}{\beta}} \prod_{i=1}^{p+1} \alpha_i^{\omega_i - 1} \tag{6}$$

The posterior function is given as

$$h_1(\underline{\alpha}|\underline{x}) = \frac{\{\prod_{t=p+1}^{T} \frac{1}{\sqrt{2\pi\sigma_t^2}} e^{-\frac{a_t^2}{2\sigma_t^2}}\} \frac{\alpha_0^{r-1} e^{-\frac{\alpha_0}{\beta}}}{\beta^r \Gamma(r)} \frac{1}{B(\omega)} \prod_{i=1}^{p} \alpha_i^{\omega_{i-1}}}{\iint \dots \int \{\prod_{t=p+1}^{T} \frac{1}{\sqrt{2\pi\sigma_t^2}} e^{-\frac{a_t^2}{2\sigma_t^2}}\} \frac{\alpha_0^{r-1} e^{-\frac{\alpha_0}{\beta}}}{\beta^r \Gamma(r)} \frac{1}{B(\omega)} \prod_{i=1}^{p} \alpha_i^{\omega_{i-1}} \, d\alpha_0 d\alpha_1 \dots d\alpha_p} \tag{7}$$

which can not be expressed in a closed form. Also, the Bayes estimator of α_i can not be written analytically. Therefore an approximation is required. As it was pointed out by Mahmoud (1991), the use of numerical computer routines may not converge for a given set of data$\underline{x}$. Thus, when the posterior is complicated, researchers have usually resorted to the Markov Chain Monte Carlo (MCMC) method and successfully achieve a satisfactory computational answer; see Robert and Casella (1999)and Li et. al. (2008).

Also, one can use Tierney's and Kadane's(1986)approximation or Lindley's method; see Lindley (1980). Nadaretc (2010) used Lindley's approximation in the Bayesian estimation of P(Y<X) for Kumaraswamy's distribution. In this study we considered Lindley's approximation for the Bayes estimation of α_i.

Lindley (1980)developed approximate procedures for the evaluation of the ratio of two integrals which are in the form of

$$\frac{\int v(\underline{\theta}) \exp\{L(\underline{\theta})\} \, d\underline{\theta}}{\int g(\underline{\theta}) \exp\{L(\underline{\theta})\} \, d\underline{\theta}} \tag{8}$$

where $\underline{\theta} = (\theta_1, \theta_2, \dots, \theta_n)$, $L(\underline{\theta})$ is the logarithm of the likelihood function, and $g(\underline{\theta})$ and $v(\underline{\theta}) = u(\underline{\theta})g(\underline{\theta})$ are arbitrary functions of $\underline{\theta}$. The posterior expectation of the function $u(\underline{\theta})$, for given $\underline{x}$, is

$$E[u(\underline{\theta})|\underline{x}] = \int u(\underline{\theta}) \exp\{L(\underline{\theta}) + \rho(\underline{\theta})\} d\underline{\theta} / \int \exp\{L(\underline{\theta}) + \rho(\underline{\theta})\} d\underline{\theta} \quad (9)$$

where$L(\underline{\theta}) + \rho(\underline{\theta})$ is the the posterior distribution of $\underline{\theta}$ except for the normalizing constant and $\rho(\underline{\theta}) = lng(\underline{\theta})$. Expanding $L(\underline{\theta}) + \rho(\underline{\theta})$ in equation (9) into a Taylor series expansion about the ML estimates of $\underline{\theta}$gives $E[u(\underline{\theta})|\underline{x}]$. So, $E[u(\underline{\theta})|\underline{x}]$ can be estimated asymptotically by

$$\hat{u}_B = u + \frac{1}{2}\sum_i \sum_j (u_{ij} + 2u_i \rho_j)\varphi_{ij} + \frac{1}{2}\sum_i \sum_j \sum_k \sum_l L_{ijk}\, \varphi_{ij}\varphi_{kl} u_l \quad (10)$$

where $i, j, m, l = 1,2, \ldots, n$, and

$$u = u(\underline{\theta}), \;\; u_i = \frac{\partial u}{\partial \theta_i}, \;\; u_{ij} = \frac{\partial^2 u}{\partial \theta_i \partial \theta_j}, \;\;\; L_{ijk} = \frac{\partial^3 L}{\partial \theta_i \partial \theta_j \partial \theta_k},$$
$$\rho_j = \frac{\partial \rho}{\partial \theta_j}, L_{ij} = \frac{\partial^2 L}{\partial \theta_i \partial \theta_j}$$

andφ_{ij} is the $(i,j)th$ element of the inverse matrix $\{-L_{ij}\}$ and all are evaluated at the MLE of the parameters.

In the next two subsections of this section Bayes estimates of the parameters of the ARCH(1) and ARCH(2) models will be derived in detail using Lindley's approximation.

2.1 Bayes Estimation of the ARCH(1) Model

For the ARCH(1) model the log likelihood reduces to

$$L = \sum_{t=2}^{T} (-\frac{1}{2}\ln(2\pi) - \frac{1}{2} lnc_t - \frac{1}{2}\frac{a_t^2}{c_t}) \quad (11)$$

where $c_t = \alpha_0 + \alpha_1 a_{t-1}^2$.

For the two parameter case: $(\theta_1, \theta_2) = (\alpha_0, \alpha_1)$. Using e.q. (10), the Bayes estimator of α_0 simplifies as

$$\alpha_{0\_SEL}^{*} = \hat{\alpha}_0 + \rho_1\varphi_{11} + \rho_2\varphi_{12} + \frac{1}{2}\{L_{111}\varphi_{11}^2 + 3L_{112}\varphi_{11}\varphi_{12}$$
$$+L_{122}(2\varphi_{12}^2 + \varphi_{22}\varphi_{11}) + L_{222}\varphi_{12}\varphi_{22}\}u_1 \quad (12)$$

where $u = a_0, u_1 = 1,\ u_2 = 0,\ \ u_{ij} = 0\ i,j = 1,2$and the Bayes estimator of α_1

$$\alpha^*_{1\_SEL} = \hat{\alpha}_1 + \rho_1\varphi_{12} + \rho_2\varphi_{22} + \frac{1}{2}\{L_{111}\varphi_{11}\varphi_{12} + L_{112}(\varphi_{11}\varphi_{22} + 2\varphi_{12}^2) + L_{122}(\varphi_{12}\varphi_{22} + 2\varphi_{12}\varphi_{22}) + L_{222}\varphi_{22}^2\}u_2 \quad (13)$$

One can obtain the values of $\{L_{ij}\}$ for $i,j = 1,2$ as

$$L_{11} = \sum_{t=2}^{T}\frac{1}{2}\left(\frac{1}{c_t^2} - \frac{a_t^2}{c_t^3}\right),\ \ L_{22} = \sum_{t=2}^{T}\frac{1}{2}\left(\frac{a_{t-1}^4}{c_t^2} - \frac{2a_t^2a_{t-1}^4}{c_t^3}\right)$$

$$L_{12} = L_{21} = \sum_{t=p+1}^{T}\frac{1}{2}\left(\frac{a_{t-1}^2}{c_t^2} - \frac{a_t^2a_{t-1}^2}{c_t^3}\right)$$

The values of $\{L_{ijk}\}$ for $i,j = 1,2,3$ are

$$L_{111} = \sum_{t=2}^{T}\left(-\frac{1}{c_t^3} + \frac{3a_t^2}{c_t^4}\right),\qquad L_{122} = L_{212} = L_{221} = \sum_{t=2}^{T}(-\frac{a_{t-1}^4}{c_t^3} + \frac{3a_t^2a_{t-1}^4}{c_t^4})$$

$$L_{112} = L_{211} = L_{121} = \sum_{t=2}^{T}(-\frac{2a_{t-1}^2}{c_t^3} + \frac{3a_t^2a_{t-1}^2}{c_t^4}),$$

$$L_{222} = \sum_{t=2}^{T} -\left(\frac{a_{t-1}^6}{c_t^3} - \frac{3a_t^2a_{t-1}^4}{c_t^4}\right) \quad (14)$$

The partial derivatives of $\rho(\underline{\theta})$ with respect to α_0 and α_1 are, $\rho_1 = (r - 1)/\alpha_0 - 1/\beta$and $\rho_2 = \frac{\omega_1 - 1}{\alpha_1} + (1 - \omega_2)/(1 - \alpha_1)$. Finally, $u_1 = 1$and $u_2 = 1$.

In the case that the prior of α_0 is not specified, one might want to use a vague prior given by $g_3(\alpha_0) \propto 1/\alpha_0$. In this situation the Bayes estimators for α_0 and α_1 denoted as $\alpha^*_{0_SEL}$and $\alpha^*_{1\_SEL}$, obtained from eq(12) and eq(13) with $\rho_1 = -{}^1\!/_{\alpha_0}$.

Instead of using the well known symmetric SE loss function, one can use the asymmetric LINEX loss function which was first proposed by Varian (1974) and is given as

$$L(\xi,\delta) = e^{\gamma(\delta-\xi)} - \gamma(\delta - \xi) - 1 \quad (15)$$

whereξ is a univariate parameter and $\gamma \neq 0$. The parameter γ is known and gives the degree of asymmetry. If γ> 0 and the errors $\delta - \xi$ are positive the LINEX loss function is almost exponential and for negative errors almost linear, in this situation overestimation is a more serious problem than underestimation. If γ< 0 underestimation is more important than overestimation. In addition to the LINEX loss function one could use other asymmetric loss functions, such as balanced LINEX, or asymmetric linear and asymmetric quadratic loss functions.

Let $M_{\Xi|\underline{x}}(t) = E_{\Theta|\underline{x}}(e^{t\xi})$ be the moment generating function of the posterior density function of Ξgiven $\underline{x}$. It can be easily verified that the value of $\delta(\Xi)$ that minimizes $E_{\Xi|\underline{x}}(L(\Xi, \delta(\Xi))$ in equation (14) is $\delta^*(\Xi) = -\frac{1}{\gamma} ln M_{\Xi|\underline{x}}(-\gamma)$, provided that $M_{\Xi|\underline{x}}(.)$ exists and is finite.

In order to compute the Bayes estimators of α_0 let $u(\alpha_0) = e^{-\gamma\alpha_0}$ then eq (10) takes the form

$$\alpha_0^{**} = e^{-\gamma\widehat{\alpha}_0} + \rho_1\varphi_{11} + \rho_2\varphi_{12} + \frac{1}{2}\{(L_{111}\varphi_{11}^2 + 3L_{112}\varphi_{11}\varphi_{12} + L_{122}(2\varphi_{12}^2 + \varphi_{22}\varphi_{11}) + L_{222}\varphi_{12}\varphi_{22})u_1 + u_{11}\varphi_{11}\} \quad (16)$$

where $u_1 = -\gamma e^{-\widehat{\gamma\alpha}_0}$, $u_{11} = \gamma^2 e^{-\gamma\widehat{\alpha}_0}$and thus the LINEX estimator of α_0 is given as

$$\alpha_{0\_LINEX}^{**} = -\frac{1}{\gamma}\ln(\alpha_0^{**}) \quad (17)$$

Similarly, in order to compute the Bayes estimators of α_1 let $u(\alpha_1) = e^{-\gamma\alpha_1}$ then the Bayes estimator of $E(e^{-\gamma\alpha_1})$ simplifies as

$$\alpha_1^{**} = e^{-\gamma\widehat{\alpha}_1} + \rho_1\varphi_{12} + \rho_2\varphi_{22} + \frac{1}{2}\{(L_{111}\varphi_{11}\varphi_{12} + L_{112}(\varphi_{11}\varphi_{22} + 2\varphi_{12}^2) + L_{122}(\varphi_{12}\varphi_{22} + 2\varphi_{12}\varphi_{22}) + L_{222}\varphi_{22}^2)u_2 + u_{22}\varphi_{22}\} \quad (18)$$

where$u_2 = -\gamma e^{-\widehat{\gamma\alpha}_1}andu_{22} = \gamma^2 e^{-\gamma\widehat{\alpha}_1}$.. Thus, the LINEX estimator of α_1 is given as

$$\alpha_{1\_LINEX}^{**} = -\frac{1}{\gamma}\ln(\alpha_1^{**}) \quad (19)$$

In the case that one wants to use a vague prior for α_0 the LINEX estimators will be given by eq(16) and eq(18) with $\rho_1 = -{}^{1}/_{\alpha_0}$.

2.2 Bayes' Estimation of Normal-ARCH(2) Model

For the ARCH(2) model the log likelihood function reduces to

$$L = \sum_{t=p+1}^{T} \left(-\frac{1}{2}\ln(2\pi) - \frac{1}{2} lnc_t - \frac{1}{2}\frac{a_t^2}{c_t}\right) \tag{20}$$

where $c_t = \alpha_0 + \alpha_1 a_{t-1}^2 + \alpha_2 a_{t-1}^2$

For the three parameter case when u is a function of only of one of the three parameters $\underline{\theta} = (\alpha_0, \alpha_1, \alpha_2)$ Lindley's approximation simplifies, when a SEL function is assumed, as

$$\alpha_{i-1\_SEL}^{**} = \hat{\alpha}_{i-1} + u_i \rho_i \varphi_{ii} + \frac{1}{2}\{Au_i\varphi_{1i} + Bu_i\varphi_{2i} + Cu_i\varphi_{3i}\} \qquad i = 1, 2, 3 \tag{21}$$

where $u_i = 1$ for $i = 1, 2, 3$, $\rho_1 = \frac{r-1}{\alpha_0} - \frac{1}{\theta}$, $\rho_2 = \frac{\omega_1 - 1}{\alpha_1} - \frac{\omega_3 - 1}{1-\alpha_1-\alpha_2}$ and $\rho_3 = \frac{\omega_2 - 1}{\alpha_2} - \frac{\omega_3 - 1}{1-\alpha_1-\alpha_2}$.

$$\begin{aligned} A &= \varphi_{11}L_{111} + 2\varphi_{12}L_{121} + 2\varphi_{13}L_{131} + 2\varphi_{23}L_{231} + \varphi_{22}L_{221} + \varphi_{33}L_{331} \\ B &= \varphi_{11}L_{112} + 2\varphi_{12}L_{122} + 2\varphi_{13}L_{132} + 2\varphi_{23}L_{232} + \varphi_{22}L_{222} + \varphi_{33}L_{332} \\ C &= \varphi_{11}L_{113} + 2\varphi_{12}L_{123} + 2\varphi_{13}L_{133} + 2\varphi_{23}L_{233} + \varphi_{22}L_{223} + \varphi_{33}L_{333} \end{aligned} \tag{22}$$

When a LINEX loss function is assumed, first we estimate $e^{-\nu\alpha_{i-1}}$ $i = 1, 2, 3$ using eq(10) which reduces to

$$\begin{aligned} (e^{-\gamma\alpha_{i-1}})_B = e^{-\widehat{\gamma\alpha}_{i-1}} + (u_1d_1 + u_2d_2 + u_3d_3 + d_4 + d_5) + \frac{1}{2}[A(u_1\varphi_{11} \\ + u_2\varphi_{12} + u_3\varphi_{13}) + B(u_1\varphi_{21} + u_2\varphi_{22} + u_3\varphi_{23}) \\ + C(u_1\varphi_{31} + u_2\varphi_{32} + u_3\varphi_{33})] \end{aligned} \tag{23}$$

where

$$\begin{aligned} d_i &= \rho_1\varphi_{i1} + \rho_2\varphi_{i2} + \rho_3\varphi_{i3}\,; \qquad i = 1,2,3 \\ d_4 &= u_{12}\varphi_{12} + u_{13}\varphi_{13} + u_{23}\varphi_{23} \\ d_5 &= 0.5(u_{11}\varphi_{11} + u_{22}\varphi_{22} + u_{33}\varphi_{33}) \end{aligned}$$

and the LINEX estimators for α_0, α_1 and α_2 are given as

$$\alpha_{i-1\_LINEX}^{\#} = -\frac{1}{\gamma}\ln((e^{-\gamma\alpha_{i-1}})_B) \tag{24}$$

The Bayes estimators for α_0, α_1 and $\alpha_2$are obtained from eq(21) through eq(23) by letting $\rho_1 = -{}^{1}/_{\alpha_0}$.

The derived $L_{ij}\, i,j = 1,2,3$ and $L_{ijk}\, i,j = 1,2,3$ and the estimated variances and covariances of the MLE are given in the followings.

$$L_{11} = \sum_{t=3}^{T} \frac{1}{2}\left(\frac{1}{c_t^2} - \frac{a_t^2}{c_t^3}\right) \qquad L_{111} = \sum_{t=3}^{T}\left(-\frac{1}{c_t^3} + \frac{3a_t^2}{c_t^4}\right)$$

$$L_{22} = \sum_{t=3}^{T} \frac{1}{2}\left(\frac{a_{t-1}^4}{c_t^2} - \frac{2a_t^2 a_{t-1}^4}{c_t^3}\right) \qquad L_{222} = \sum_{t=3}^{T} -\left(\frac{a_{t-1}^6}{c_t^3} - \frac{3a_t^2 a_{t-1}^6}{c_t^4}\right)$$

$$L_{33} = \sum_{t=3}^{T} \frac{1}{2}\left(\frac{a_{t-2}^4}{c_t^2} - \frac{2a_t^2 a_{t-2}^4}{c_t^3}\right) \qquad L_{333} = \sum_{t=3}^{T} -\left(\frac{a_{t-2}^4}{c_t^3} - \frac{3a_t^2 a_{t-2}^6}{c_t^4}\right)$$

$$L_{12} = \sum_{t=3}^{T} \frac{1}{2}\left(\frac{a_{t-1}^2}{c_t^2} - \frac{a_t^2 a_{t-1}^2}{c_t^3}\right) \qquad L_{122} = \sum_{t=3}^{T} (-\frac{a_{t-1}^4}{c_t^3} + \frac{3a_t^2 a_{t-1}^4}{c_t^4})$$

where $L_{12} = L_{21}$ where $L_{122} = L_{212} = L_{221}$

$$L_{13} = \sum_{t=3}^{T} \frac{1}{2}\left(\frac{a_{t-2}^2}{c_t^2} - \frac{a_t^2 a_{t-2}^2}{c_t^3}\right) \qquad L_{133} = \sum_{t=3}^{T} \frac{1}{2}\left(-\frac{a_{t-2}^4}{c_t^3} + \frac{a_t^2 a_{t-2}^4}{c_t^4}\right)$$

where $L_{13} = L_{31}$ where $L_{133} = L_{313} = L_{331}$

$$L_{23} = \sum_{t=3}^{T} \frac{1}{2}\left(\frac{a_{t-1}^2 a_{t-2}^2}{c_t^2} - \frac{a_t^2 a_{t-1}^2 a_{t-2}^2}{c_t^3}\right) \qquad L_{233} = \sum_{t=3}^{T} (-\frac{a_{t-1}^2 a_{t-2}^4}{c_t^3} + \frac{3a_t^2 a_{t-1}^2 a_{t-2}^4}{c_t^4})$$

where $L_{23} = L_{32}$ where $L_{233} = L_{332} = L_{323}$

$$L_{112} = \sum_{t=3}^{T} (-\frac{2a_{t-1}^2}{c_t^3} + \frac{3a_t^2 a_{t-1}^2}{c_t^4}) \qquad L_{223} = \sum_{t=3}^{T} (-\frac{a_{t-1}^4 a_{t-2}^2}{c_t^3} + \frac{3a_t^2 a_{t-1}^4 a_{t-2}^2}{c_t^4})$$

where $L_{112} = L_{121} = L_{211}$ where $L_{223} = L_{232} = L_{322}$

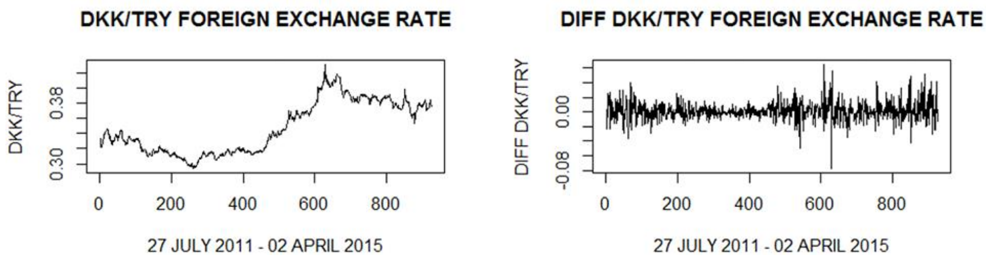
$$L_{113} = \sum_{t=3}^{T} (-\frac{2a_{t-2}^2}{c_t^3} + \frac{3a_t^2 a_{t-2}^2}{c_t^4}) \qquad L_{123} = \sum_{t=3}^{T} (-\frac{a_{t-1}^2 a_{t-2}^2}{c_t^3} + \frac{3a_t^2 a_{t-1}^2 a_{t-2}^2}{c_t^4})$$

where $L_{113} = L_{131} = L_{311}$ where $L_{123} = L_{321} = L_{213} = L_{132} = L_{231} = L_{312}$

3. Example

In order to illustrate the findings of this study an example with real data is given. The data are taken from the Turkish Central Bank which shows daily foreign exchange rate of DKK (Danish Korona) versus TRY (Turkish Lira) for the time period from July 2011 to April 2015. Dickey-Fuller and Phillips-Perron unit root tests show that the data are not stationary. By taking the first difference the series becomes stationary. The graphs of the original and first difference series are shown in Figure 1.

Figure 1.Time series plot and first difference of DKK/TRY



The Lagrange Multiplier (LM) test for ARCH effects is performed and with p-value < 2.2e-16 we can conclude that an ARCH model can be used to fit the data. For illustrative purposes an ARCH(1) model will be utilized with standard normal innovations. The ML and Bayes estimates, under SE and LINEX loss functions when $\gamma = 0.3$ of the parameters are computed. For the parameter α_0 a vague prior is assumed whereas the parameters ω_1 and ω_2 of the beta prior of α_1 are estimated from the data by equating the sample mean and variance with the population mean and variance. The resulting estimates are $\widehat{\omega}_1$=0.0003782319 and $\widehat{\omega}_2$=8.397714. In addition of estimating the parameters the 1115 values of the differenced series will be used to estimate the parameters and then using the ARCH(1) model the next 10 values will be predicted. The estimated predicted values will be compared to the real ones by computing the mean errors. Table 1 below shows the estimates of the parameters when the innovations are follow the standard normal distribution.

Table 1.Parameter estimates for the ARCH(1) with normal innovations

Coefficients	MLEs	Bayes SE	Bayes LINEX
α_0	2.864901e-06	2.848307e-06	2.848307e-06
α_1	0.4286469	0.4341838	0.4342412

Forecasting with the ARCH(p) model $\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \alpha_2 a_{t-2}^2 + \dots + \alpha_p a_{t-p}^2$ can be obtained recursively. At the forecast origin h, the 1-step ahead forecast of σ_{h+1}^2 is

$$\sigma_h^2(1) = \alpha_0 + \alpha_1 a_h^2 + \alpha_2 a_{h-1}^2 + \dots + \alpha_p a_{h+1-p}^2$$

The 2-step ahead forecast is

$$\sigma_h^2(2) = \alpha_0 + \alpha_1 a_h^2(1) + \alpha_2 a_h^2 + \dots + \alpha_p a_{h+2-p}^2$$

and the l- step ahead forecast can be found as following

$$\sigma_h^2(l) = \alpha_0 + \sum_{i=1}^{p} \alpha_i a_h^2(l-i)$$

where $\sigma_h^2(l-i) = a_{h+l-i}^2$ if $(l-i) < 0$.

The ML and Bayes estimates of $\sigma_h^2(l)$ are derived by replacing the parameters $\alpha_0, \alpha_1, .., \alpha_{p-1}$ and α_p by the corresponding estimates.

Table 2 shows the real data along with the estimated forecasted values when the innovations are normally distributed with mean zero and standard deviation 1. Table 3 shows the mean errors.

Table 2. Actual values and estimates of $\sigma_h^2(1)$ using an ARCH(1) model with normal innovations

ACTUAL	MLE	BAYES (SE)	BAYES(LINEX)
0.00071	0.00175355	0.001749588	0.001749596
0.00072	0.002045229	0.002043861	0.002043907
0.00143	0.00215822	0.002159179	0.002159253
0.00144	0.002204881	0.002207373	0.002207465
0.00031	0.002224583	0.002227973	0.002228076
0.0073	0.002232975	0.002236859	0.002236967
0.00394	0.002236562	0.002240706	0.002240816
0.00012	0.002238098	0.002242374	0.002242486
0.00316	0.002238756	0.002243098	0.002243211
0.00162	0.002239038	0.002243412	0.002243525

Table 3.Mean Square Error of Forecasts

SE loss function		LINEX loss function	
ML	**Bayes**	**ML**	**Bayes**
4.192089e-06	4.188936e-06	1.886957e-07	1.885523e-07

4. Simulation results

The Bayesian and ML estimators were compared by means of Monte Carlo simulations. The simulation study is undertaken using a standard normal or a standardized student-t distribution for the innovations and different sample sizes. In particular the sample sizes are 200, 400, 600, 800 and 1000. The prior for α_0 is a gamma or an improper prior and for $(\alpha_{1,}\alpha_j)$ $j = 1{,}2$ Dirichlet function. Using the above mentioned innovations, sample sizes and priors we obtain the ML and Bayes estimates of the parameters under a SE and LINEX loss functions. Tables 1 through 4 present the mean true value for each parameter and the average values of the ML and Bayesian estimates are reported. Furthermore, in parenthesis, the mean errors are also reported. All the results are based on 1000 repetitions. In all simulations the value of the LINEX parameter γ is -0.5.

In particular, in Table 4 considers the ARCH(1) model when the innovations are drawn from the standard normal distribution. The prior of α_0 is either a gamma with $r = 2$ and$\beta = 1$ or a vague prior and the Dirichlet prior of α_1 is beta with parameters $\omega_1 = 1$ and $\omega_2 = 3$.

Table 5 is similar to table one, but considers an ARCH(2) model with standard normal innovations. The prior of α_0 is either a gamma with $r = 3$ and $\beta = 1$ or a vague prior and a Dirichlet prior for α_1, α_2 with parameters $\omega_1 = 1$,$\omega_2 = 3$ and $\omega_3 = 2$.

From Tables 4 and 5 it is observed that as the sample sizes increase the MSEs and MEs decrease. This should be expected since the MLEs are consistent. Also, as expected for all the estimates when the sample sizes increase the MSEs and MEs decrease. In all cases the MSEs and MEs when proper priors are used for the Bayes estimates are smaller than the ones corresponding when an improper prior is used for α_0 and to the MLE estimates. Finally, there is little difference between the MLEs and MEs when an improper prior for α_0 is utilized.

Table 4.ARCH(1) model with standardized normal innovations

AVERAGE VALUES					
α_0 Gamma(2,1)	1.99592	1.99710	1.97473	2.04904	1.99442
α_1 Beta(1,3)	0.28128	0.26563	0.26527	0.258863	0.264296
		Sample	**Size**		
SE LOSS FUNCTION	**200**	**400**	**600**	**800**	**1000**
MLEs of α_0	2.08847	2.06571	2.02353	2.10426	2.02547
	(0.16551)	(0.06509)	(0.04130)	(0.03928)	(0.02875)
α_0 (vague)	1.99614	2.02881	2.00171	2.08832	2.01417
	(0.14502)	(0.05558)	(0.03524)	(0.03487)	(0.02645)
α_0(Gamma)	1.95067	2.00670	1.98795	2.07698	2.00523
	(0.13911)	(0.05091)	(0.03238)	(0.03076)	(0.02438)
MLEs of α_1	0.28128	0.26563	0.26527	0.25886	0.26430
	(0.01684)	(0.00763)	(0.00542)	(0.00390)	(0.00304)
α_1 (Beta), α_0(vague)	0.29681	0.26316	0.26067	0.25121	0.25829
	(0.01418)	(0.00644)	(0.00432)	(0.00331)	(0.00263)
α_1 (Beta), α_0(Gamma)	0.29735	0.26341	0.26078	0.25145	0.25832
	(0.01400)	(0.00636)	(0.00429)	(0.00328)	(0.00263)
LINEX LOSS FUNCTION					
MLEs of α_0	2.08847	2.06571	2.02353	2.10426	2.02547
	(0.02314)	(0.00854)	(0.00542)	(0.00518)	(0.00371)
α_0(vague)	2.10712	2.07565	2.03058	2.10933	2.02969
	(0.02215)	(0.00836)	(0.00524)	(0.00502)	(0.00351)
α_0(Gamma)	2.12557	2.08485	2.03688	2.11366	2.03317
	(0.02032)	(0.00820)	(0.00514)	(0.00492)	(0.00344)
MLEs of α_1	0.28128	0.26563	0.26527	0.25886	0.26430
	(0.00208)	(0.00095)	(0.00067)	(0.00048)	(0.00038)
α_1(Beta), α_0(vague)	0.28950	0.26582	0.26493	0.25678	0.26301
	(0.00207)	(0.00092)	(0.00066)	(0.00047)	(0.00037)
α_1(Beta), α_0(Gamma)	0.29662	0.26529	0.26351	0.25359	0.26065
	(0.00190)	(0.00090)	(0.00065)	(0.00046)	(0.00036)

Table 5. ARCH(2) model with standardized normal innovations

AVERAGE VALUES					
α_0 Gamma(3,1)	2.86877	2.88818	2.98862	2.91820	2.95684
α_1 Dirichlet(1,3,2)	0.20130	0.19375	0.19138	0.18543	0.18391
α_2 Dirichlet(1,3,2)	0.35504	0.34376	0.34070	0.33717	0.33700
		Sample	Size		
SE LOSS FUNCTION	200	400	600	800	1000
MLEs of α_0	3.20131	3.13150	3.20031	3.09812	3.03502
	(0.63128)	(0.32351)	(0.21924)	(0.13268)	(0.09963)
α_0(vague)	3.04339	3.05814	3.15375	3.06203	3.00845
	(0.45898)	(0.26875)	(0.19114)	(0.11548)	(0.09473)
α_0 (Gamma)	2.98106	3.03352	3.13576	3.05021	3.00029
	(0.35373)	(0.23300)	(0.17011)	(0.10354)	(0.09185)
MLEs of α_1	0.18202	0.17216	0.17290	0.16913	0.17286
	(0.01088)	(0.00537)	(0.00350)	(0.00189)	(0.00186)
α_1 (Dirichlet), α_0 (vague)	0.19852	0.18044	0.17806	0.17323	0.17611
	(0.01012)	(0.00522)	(0.00326)	(0.00171)	(0.00178)
α_1 (Dirichlet) , α_0(Gamma)	0.19896	0.18055	0.17815	0.17327	0.17612
	(0.00997)	(0.00518)	(0.00325)	(0.00170)	(0.00178)
MLEs of α_2	0.29803	0.30506	0.30234	0.30535	0.31690
	(0.01703)	(0.00822)	(0.00582)	(0.00311)	(0.00303)
α_2 (Dirichlet), α_0(vague)	0.31335	0.31213	0.30669	0.30841	0.31919
	(0.01468)	(0.00766)	(0.00544)	(0.00292)	(0.00290)
α_2(Dirichlet) , α_0(Gamma)	0.31384	0.31227	0.30680	0.30845	0.31921
	(0.01456)	(0.00763)	(0.00541)	(0.00291)	(0.00290)
LINEX LOSS FUNCTION					
MLEs of α_0	3.20131	3.13150	3.20031	3.09812	3.03502
	(0.07891)	(0.04044)	(0.02740)	(0.01659)	(0.01245)
α_0(vague)	3.22268	3.19514	3.17735	3.18040	3.03206
	(0.05439)	(0.03490)	(0.02450)	(0.01563)	(0.01225)
α_0(Gamma)	3.21321	3.16454	3.19005	3.14048	3.03476
	(0.04422)	(0.02913)	(0.02126)	(0.01294)	(0.01148)
MLEs of α_1	0.18202	0.17216	0.17290	0.16913	0.17286
	(0.00181)	(0.00089)	(0.00058)	(0.00031)	(0.00031)
α_1 (Dirichlet), α_0(vague)	0.19141	0.17744	0.17662	0.17232	0.17562
	(0.00177)	(0.00087)	(0.00056)	(0.00029)	(0.00030)
α_1 (Dirichlet) , α_0(Gamma)	0.18781	0.17589	0.17785	0.17182	0.17533
	(0.00166)	(0.00086)	(0.00054)	(0.00028)	(0.00030)
MLEs of α_2	0.29803	0.30506	0.30234	0.30535	0.31690
	(0.00310)	(0.00149)	(0.00106)	(0.00057)	(0.00055)
α_2 (Dirichlet),α_0(vague)	0.31251	0.30182	0.30566	0.30635	0.31812
	(0.00290)	(0.00143)	(0.00099)	(0.00053)	(0.00053)
α_2(Dirichlet) , α_0(Gamma)	0.30513	0.30330	0.30387	0.30571	0.31737
	(0.00265)	(0.00139)	(0.00098)	(0.00053)	(0.00053)

5. Conclusions

In this study, we have considered the Bayesian inference of ARCH(p) model with normal distributed innovations. The Lindley's approximation method is applied to obtain the Bayesian estimator of unknown parameters since the Lindley's approximation is a one of the proper methods in case which the Bayes estimators of ARCH model cannot be obtained in explicit forms. ARCH(1) and ARCH(2) models are studied due to the fact that Lindley method is difficult to compute when the number of unknown parameters is increased. In this situation, one can use Tierney-Kadane method yet this method requires two maximizations. Furthermore, the Lindley method is as accurate as Tierney-Kadane method since sample sizes are enough large in the simulation study.

At first hand, it is observed that as expected, the MSE of each estimator decreases as the sample size increases. Bayes estimators obtained from Lindley's method are observed to perform much better than the MLEs under SEL and LINEX functions, but the discrepancy in their relative performance tends to get smaller and smaller with the increase in sample size because of the consistency of ML method.

We have also considered the forecasting performance of Bayes estimates and MLEs using foreign exchange data as an example. Although Bayes estimates are not always better than MLEs, it is observed that Bayes estimator is quite well in prediction especially predictors of ARCH(1) model with normal innovations.

REFERENCES

[1]**Batchelor R. and Peel D.A. (1998)*, Rationality Testing under Asymmetric Loss;*** *Economics Letters 61, pp. 49-54;*

[2]**Bollerslev T. and Engle R.F. and Nelson D.B. (1994), *ARCH Models*;** *In: Engle, R.F., McFadden, D. (Eds.), Handbook of Econometrics, Vol. IV. North-Holland, Amsterdam, pp. 2959-3038;*

[3]**Bose A. and Mukherjee K. (2003)*, Estimating the ARCH Parameters by Solving Linear Equations;*** *Journal of Time Series Analysis, 24(2), pp. 127–136.*

[4] **Christian Robert and George Casella (1999)*, Monte Carlo Statistical Methods;*** *Springer-Verlag,NewYork;*

[5]**ChristoffersenP. F. and Diebold F. X. (1997), *Optimal Prediction under Asymmetric Loss;*** *Econometric Theory, Cambridge University Press, vol. 13-06, pp. 808-817;*

[6]**Clive William J Granger; Paul Newbold(1986)*, Forecasting Economic Time Series;*** *Orlando: Academic Press;*

[7]**DanielStraumann(2005*, Estimation in Conditionally Heteroscedastic Time Series Models, Lecture Notes in Statistics;*** *Berlin: Springer-Verlag, 181;*

[8]**Engle, R. F. (1982),** ***Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation;****Econometrica, 50 (4), pp. 987–1007;*
[9] **George E. P. Box and George C. Tiao (1992),** ***Bayesian Inference in Statistical Analysis;*** *Wiley and Sons;*
[10]**Gilks W.R. and Best,N.G.andTan,K.K.C.(1995),** ***Adaptive Rejection Metropolis Sampling within Gibbs Sampling;****Journal of the Royal Statistical Society Series C 44: 455–472;*
[11]**Giraitis L. and Robinson P. M. (2001),** ***Whittle Estimation of ARCH Models;****Econometric Theory, 17(3), pp. 608–631;*
[12]**Granger C. W. J.and Pesaran M. H. (2000),** ***Economic and Statistical Measures of Forecast Accuracy;*** *J. Forecast., 19, pp. 537–560;*
[13]**Kim S. and Shephard, N. and Chib, S. (1998),** ***Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models****; Review of Economic Studies 65: 361–393;*
[14]**Li J., Zhang C.M., Nordheim E.V. andLehner C.E. (2008);** ***On the Multivariate Predictive Distribution of Multi-dimensional Effective Dose: A Bayesian Approach;*** *J. Statist. Comput.Simul. 78(5) pp. 429–442;*
[15]**Lindley D.V. (1980),** ***Approximate Bayes Methods*****;** *Trabajos de Estadistica 3 pp. 281–288;*
[16]**Mahmoud M. (1991),** ***Bayesian Estimation of the 3-parameter Inverse Gaussian Distribution;*** *Trabajos de Estadistica 6 pp. 45–62;*
[17]**Metropolis N. and Rosenbluthv A. W. and Rosenbluth M. N. and Teller A. H. and Teller E. (1953),** ***Equation of State Calculations by Fast Computing Machines;*** *J. Chem. Phys. 21, pp. 1087-1092;*
[18]**Moore D. and Papadopoulos A.S. (2000),** ***The Burr Type XII distribution as a Failure Model under Various Loss Functions;*** *Microelectronics Reliability, 40, pp. 2117-2122.*
[19]**NadarM.,Papadopoulos A. and Kızılaslan F. (2012),** ***Classical and Bayesian Estimation of P(Y < X) for Kumaraswamy's Distribution;****Journal of Statistical Computation and Simulation, 84:7, pp. 1505-1529;*
[20]**Peel D.A. and Nobay A.R. (1998),** ***Optimal Monetary Policy in a Model of Asymmetric Central Bank Preferences;*** *LSE - FMG discussion paper 0306;*
[21] **Rue S. Tsay (2012),** ***Analysis of Financial Time Series;*** *Wiley and Sons;*
[22]**Tierney L. and Kadane J.B. (1986),** ***Accurate Approximations for Posterior Moments and Marginal densities;****J. Am. Statist. Assoc. 81 pp.82–86;*
[23]**Varian H. R. (1974),** ***A Bayesian Approach to Real Estate Assessment;****In: Studies in Bayesian econometrics and statistics in honor of Leonard J, Savage (ed. S.E. Fienberg and A. Zellner) North Holland, pp. 195-208;*

[24]**West K.D.,Edison H.J. and Cho D. (1993*), A Utility-based Comparison of Some Models of Exchange Rate Volatility;*** *Journal of International Economics 35, 23-46;*
[25]**Zellner A. (1986),** ***Bayesian Estimation and Prediction Using Asymmetric Loss Functions;*** *Journal of the American Statistical Association, 81, pp. 446-451.*