Original article

Estimating the probability distribution of the exchange rate between Ghana Cedi and American dollar

Solomon Sarpong

Department of Statistics, Faculty of Mathematical Sciences, University for Development Studies, Navrongo Campus, Ghana

ABSTRACT

The aim of this research is to find the best probability distribution function (PDF) that fits the data on exchange rate between the Ghana Cedi and American dollar. Economic data in recent times is of much importance to many persons as well as the managers of a country as a whole. Volatilities in these data to a large extent have significant effect on the individuals and the country. The interest in exchange rate volatility by researchers can be attributed to the fact that, it is empirically difficult to predict future exchange rates precisely. The volatility of the US dollar/Ghana Cedi exchange rate is high. This has necessitated the need to have a reliable method for evaluating and minimizing the risk of trading in these currencies. Knowing the type of distribution the exchange rate between the two currencies follow will give stakeholders adequate information to prevent future shocks and loses from any volatility. The data used in this research was the monthly exchange rate between the Ghana Cedi and the American dollar from the year 2000 to 2017. From the analysis, it was observed that the data follow the lognormal distribution. Hence, this analysis has shown that the exchange rate between the Ghana Cedi and the US dollar is best modeled by the lognormal distribution. The Lognormal distribution gives a very good fit to the distribution of the Ghana Cedi and the American dollar.

© 2018 The Author. Production and hosting by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Foreign exchange rate between two currencies specifies how much one currency is worth in terms of the other. A correct or appropriate exchange rate has been one of the most important factors for the growth in the economy of many countries. Exchange rate has direct influence on employment, trade flow, balance of payment, the arrangement of production and consumption, as well as outside loans of a country, Aron et al. (1998).

The world is now a global village hence, the events at one end of the globe affects the others at the extreme end. Individuals and nations trade among themselves. Since the various countries have their own currencies, trading is done in a common currency to facilitate the smoothness of the trade. Since 1944 (Bretton Woods Agreement), the American dollar has been accepted as the currency for international trade. Ghana for that matter also trades with other countries using the American dollar. There have been a lot of volatilities in the exchange rate between the Ghana Cedi and the American dollar. These volatilities have dire consequences on the economy of Ghana. The interest in exchange rate volatility by researchers can be attributed to the fact that, it is empirically difficult to predict future exchange rates precisely.

Many researchers use Box-Jenkins methodology of time series for modeling and forecast. Using Box-Jenkins methodology for forecasting has its deficiencies. This methodology assumes that there is the existence of linear relationship between the variables. But in real-world, time series data are often nonlinear, Lin et al. (2012), Huang et al. (2010), Ding et al. (2009) and Gradojevic and Yang (2006). Secondly, using the Box-Jenkins methodology for model selection procedure depends greatly on the competence and experience of the researchers, Ridhwan et al., 2015.

Hence, using the Box-Jenkins methodology for modeling and forecasting does not give enough accuracy. It is in lieu of this, this paper investigated the type of PDF that the data follows. PDFs are based on the actual data, hence are forward-looking. Also, they do not require a long historical time series in order to be estimated accurately, and furthermore are instantly capable of reflecting a change in the data. Furthermore, PDFs are well-suited for capturing the uncertainty inherent in the data. PDFs are relatively free of mathematical priors Chang and Melick (1999).

Using high-frequency data on Deutschemark and Yen returns against the dollar, Andersen et al., 2001 constructed model-free estimates of daily exchange rate volatility and correlation that cover an entire decade. Their observations included simple normality-inducing volatility transformation, high contemporaneous correlation across volatilities, high correlation between correlation and volatilities, pronounced and persistent dynamics in volatilities and correlations, evidence of long-memory dynamics in volatilities and correlations, and remarkably precise scaling laws under temporal aggregation.

Carroll (2014) found that taking a Bayesian approach, using Markov Chain Monte Carlo (MCMC) and relying on the well-known close relationship between Bayesian posterior means and maximum likelihood, the latter not computationally feasible was able to develop standard errors using balanced repeated replication, a survey-sampling approach. Fujihara and Park (1990) observed that, foreign exchange rate that are unconditional distribution of future prices change measured at high frequency (daily/weekly) has greater area in the tail relative to the normal distribution. However, at a relatively low frequency (monthly/quarterly), the unconditional distribution appears to fit the normal distribution fairly well.

Akgiray and Booth (1987) in their research on the compound distribution models of stock returns: an empirical comparison observed that the distribution of monthly returns can be safely assumed to be approximately normally distributed. Empirical studies by Tucker and Pond (1988) noted that, distribution of exchange returns are long tailed and leptokurtic to satisfy normality. Four processes were investigated in order to find their potential model. They observed discontinuity in the exchange rate and non-stationary sample in their economic appeal. The result favored mixed jump model for all six major trading currencies tested.

In an attempt to fit different PDFs to the data on the daily percentage returns of the Standard & Poor's (S&P 500), the t-distribution with location/scale parameters was shown to be an excellent choice, Egan (2007). The assumption that daily stock returns are normally distributed has long been disputed, Aparicio and Estrada, 1996. Daily stock returns typically have non-normal and asymmetric distributions (Fama, 1976; Brooks, 2002) no matter how large the data is (Amado, 2002; Brooks, 2002; Bai et al., 2003; Corrado and Zivney, 1992). From the research work done on exchange rate between different currencies, it has been observed that the distribution of exchange rate is non-normal and asymmetric. Hence, this research seeks to find the PDF that the exchange rate between the American dollar and Ghana Cedi follow. With the PDF known, more accurate and adequate forecasting can be done.

2. Materials and methods

Some PDFs will be reviewed in order to test the one that best fits the data on exchange rate between the American dollar and the Ghana Cedi.

2.1. Weibull distribution function

The random variable X is a Weibull distribution with parameters $\alpha > 0, \beta > 0$ if the pdf of X is given by

$$f(X, \alpha, \beta) = \begin{cases} \frac{\alpha}{\beta^\alpha} x^e - 1\alpha, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{1}$$

This distribution has an expectation of $E(X) = \beta(1 + 1\alpha)$ and a variance $Var(X) = \beta^2 \left\{ (1 + 2\alpha - [(1 + 1\alpha)^2]) \right\}$.

2.2. Gamma distribution function

A continuous random variable X has the standard Gamma distribution function with parameters $\alpha > 0$ and $\lambda > 0$ if its probability distribution function is given by

$$f(X) = \begin{cases} \frac{1}{(\alpha)} \lambda^\alpha x^{\alpha-1} e^{-\lambda x}, & x > 0 \\ 0, & otherwise \end{cases}.$$ The Gamma distribution function has an expectation, $E(X) = \lambda$ and variance, $Var(X) = \lambda$.

2.3. Exponential distribution function

The random variable x is an exponential probability distribution function with location and scale parameter μ and β respectively if $f(X) = \frac{1}{\beta} e^{\frac{-(x-\mu)}{\beta}, x \geq \mu; \beta > 0}$. The exponential distribution function has an expectation, $E(X) = \beta$ and variance, $Var(X) = \beta$.

2.4. Lognormal distribution function

A non-negative random variable X is said to have a lognormal distribution if the random variable $y = ln(x)$ has a normal distribution. The resulting PDF of the lognormal random variable when $ln(x)$ is normally distributed with parameters μ and σ is given by;

$$f(X; \mu, \sigma) = \begin{cases} 1\sqrt{2\pi}\sigma x \sqrt{2\pi}\sigma x e^{-\frac{|ln(x)-\mu|^2}{2\sigma^2}}, & x \leq 0 \\ 0, & x < 0 \end{cases}$$

The probability distribution function curve is as shown in Fig. 4. The expectation and the variance of the Lognormal distribution are given by; $E(X) = e^{(\mu + \sigma^2 2/2)}$ and $Var(X) = e^{[2(\mu + \sigma^2)]} - e^{(2\mu + \sigma^2)}$ respectively.

2.5. Akaike and Bayesian information criteria

These information criteria were used to compare the distribution that best fits the data. Even though both Akaike and Bayesian information criteria helps in obtaining the best model, each of these try to balance model fit and parsimony of variables and penalizes differently for the number of parameters. In order to get the best model fit, both statistics will be used.

2.6. Measuring information lost using Kullback-Leibler divergence

In quantifying how much information is in a data, the entropy for the probability distribution is calculated. The entropy probability distribution is given by

$$H = -\sum_{i=1}^{N} p(x_i) log p(x_i)$$

By modifying this equation gives the Kullback-Leibler divergence. The Kullback-Leibler divergence gives exactly how much information is lost when a probability distribution is approximated by another. Kullback-Leibler divergence is given by

$$D_{KL}(p||q) = \sum_{i=1}^{N} p(x_i) log \frac{p(x_i)}{q(x_i)}$$ where $p(x_i)$ and $q(x_i)$ are the real and approximate probability distributions respectively. The smaller Kullback-Leibler divergence value, the better it is. A smaller Kullback-Leibler divergence value is, the less information is lost in using that probability distribution as an approximation to a data.

2.7. Gaussian mixture models

This is a probability distribution that consists multiple probability distributions. For d dimensional Gaussian distribution of a vector $X = (x^1, x^2, \ldots, x^d)^T$ is defined by $N(x|\mu, \Sigma) = \frac{1}{(2\pi)^{d/2}\sqrt{|\Sigma|}} exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right)$; where μ is the mean and Σ

is the variance-covariance matrix. Gaussian Mixture Models tend to find k clusters in a data by minimizing $\frac{(x-\mu)^2}{\sigma^2}$.

3. Results and discussions

The data was the monthly exchange rate of the Ghana Cedi and the American dollar from the year 2000 to 2017. However, only the data from 2000 to 2015 was used as the training data and data from 2016 to 2017 was used for the validation of the model. During the period understudy, the exchange rate between GHS 0.99–GHS 6.67 to the dollar with a mean of 2.95 (standard error, 0.10).

The quest for this paper is to find a PDF that best approximates the data. In order to do this, a plot of the empirical density and the cumulative distributions of the data were made. From the empirical probability density function in Fig. 1, it was observed that the graph tails off to the right. Hence, the data is positively skewed.

The bootstrapping technique was used in our quest to obtain the Cullen and Frey graph, Fig. 2. From this Figure it was observed that, several distributions were possible fit for the data. From the summary statistics in the Cullen and Frey graph, it was observed that the skewness was positive, 1.23 and kurtosis was 4.81. Hence, the right tail of the data is heavier than the left tail and the data is leptokurtic. These statistics indicate that the possible PDF is positively skewed. As shown in Fig. 2, the distributions in the data include: Weibull, Gamma, Exponential, Logistic and Lognormal, normal, and Beta. However, Fig. 3 shows the distributions with positive skewness. Hence, only the Weibull, Gamma, Exponential and the lognormal distributions will be investigated in order to verify if the data follows any of these distributions.

Fig. 4 shows the theoritical Q-Q plots for the various positively skewed PDFs under consideration. This plot indicates which distribution function the data is more likely to follow. It can be seen that, Gamma distribution deviates remarkable from the hypothetical distribution. Hence, the likely distribution can either be a Weibull, lognormal or exponential distribution. This is also confirmed in the empirical and theoretical cumulative distribution function plot in Fig. 5 and P-P plot in Fig. 6.

The goodness-of-fit criteria were used to ascertain the best distribution that fits the data. The statistics that were used include log-likelihood, Bayesian and Akaike's information criteria. The best

distribution is the one that has the lowest values for these statistics. From Table 1, it can be observed that lognormal distribution has the lowest values for all these statistics. Hence, the data is likely to be lognormally distributed.

The goodness-of-fit statistics were used to test whether the data is actually lognormally distributed. The Kolmogorov-Smirnov, Cramer-von Mises and Anderson-Darling statistics were used. From the analysis and based on these statistics, it was observed that the data follows the Lognormal distribution. Hence, it has been shown that the exchange rate between the Ghana Cedi and the American dollar follows the Lognormal distribution. It was observed that, this Lognormal distribution has a log-mean value of 0.94 and log-standard deviation of 0.48.

In order to confirm that our data follows the Lognormal distribution, data was generated from simulated Lognormal distribution. The simulated Lognormal distribution has a logmean value
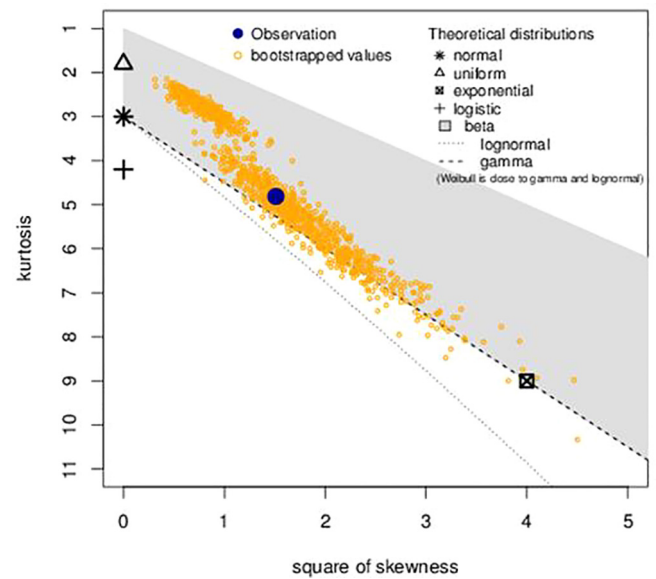
Fig. 2. Skewness and Kurtosis plot for the exchange rate between the Ghana Cedi and the American dollar.
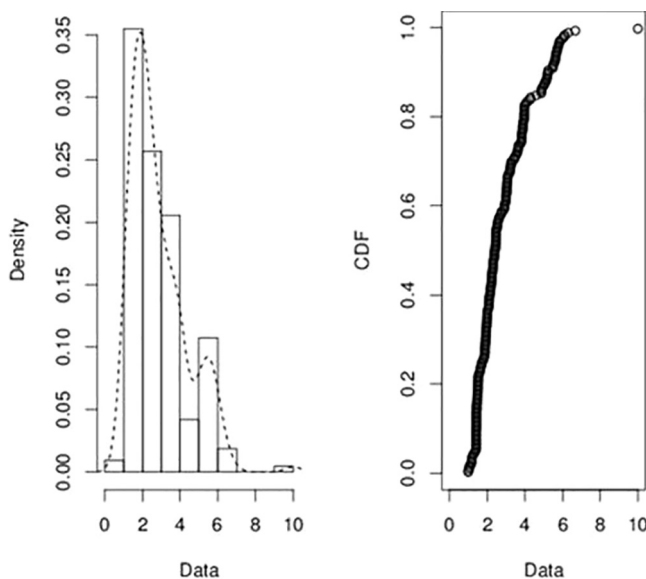
Fig. 1. The empirical density (left) and Cumulative distribution function (right) plots for the exchange rate between the Ghana Cedi and the American dollar.
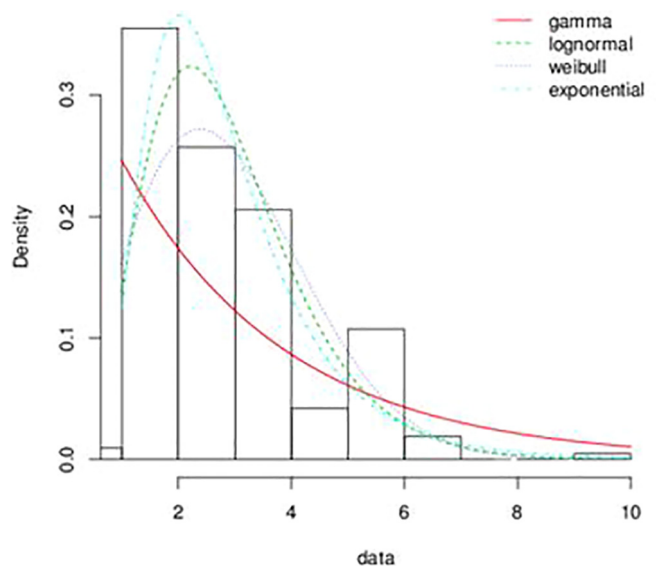
Fig. 3. A plot of the histogram and theoretical densities for the exchange rate between the Ghana Cedi and the American dollar.
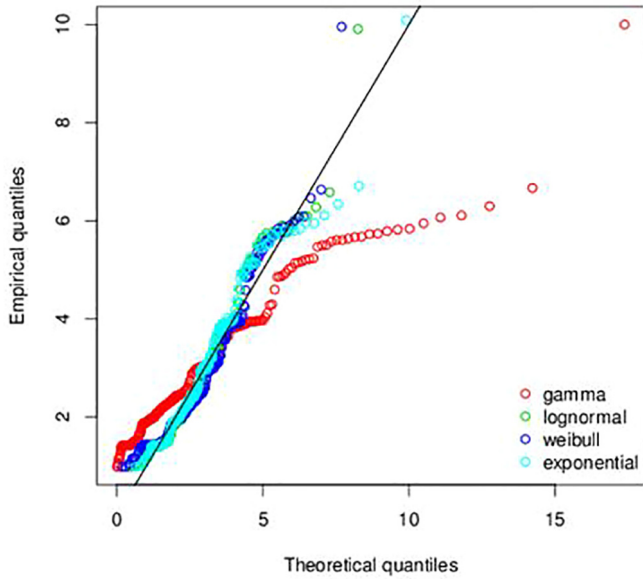
Fig. 4. QQ-plot for the exchange rate between the Ghana Cedi and the American dollar.
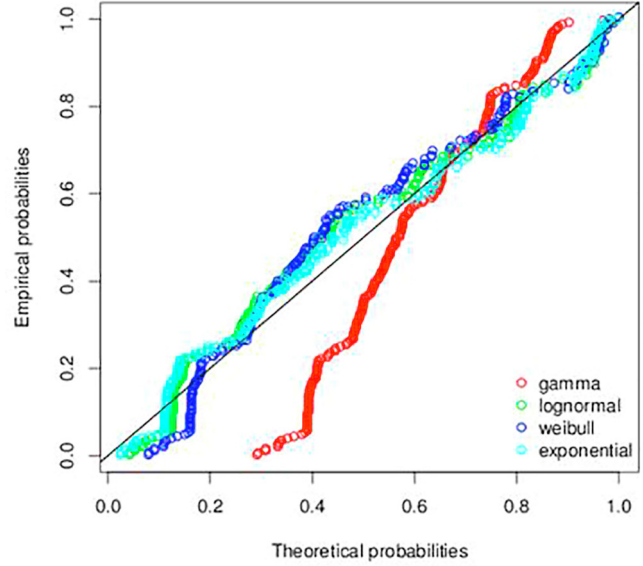

Fig. 6. P-P plot of the exchange rate between the Ghana Cedi and the American dollar.
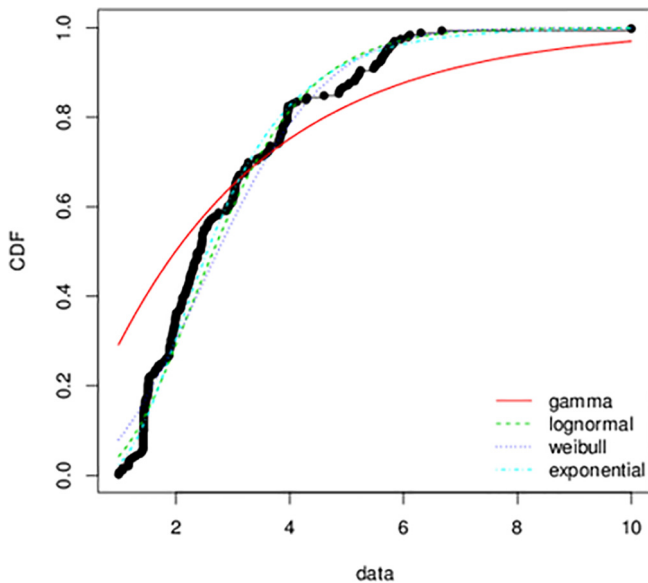

Fig. 5. Empirical and theoretical cumulative distributions functions for the exchange rate between the Ghana Cedi and the American dollar.

Table 1
The Goodness-of-fit Criteria for the exchange rate between the Ghana Cedi and the American dollar.

Probability Distribution	Akaike's Information Criterion	Bayesian Information Criterion
Weibull	735.5650	742.3284
Gamma	711.2458	717.9778
Lognormal	697.5699	704.3018
Exponential	880.9186	884.2846

or not. From the test, it was observed that there was a *p*-value of 0.4132. Hence with a significance level of 0.05, it can be observed that the two datasets are identical. It can therefore be concluded that the original data on the exchange rate between the Ghana Cedi and the American dollar follows the Lognormal distribution.

One way of judging a derived probability density function is its empirical performance – i.e. its ability to predict accurately, Chang and Melick (1999). Hence, in order to verify the practicability of using the lognormal distribution for forecasting, how best the simulated data approximates the original data was investigated. In order to do this, the Lognormal distribution was used to forecast the exchange rate between the Ghana Cedi and American dollar from 2016 to 2017. The forecast data was compared with the remaining of the original data (2016–2017). The comparison was done using two sample *t*-test. From the analysis, the *t*-value was −32.42, *p*-value = 0.12. Hence, it can be seen that there is no difference in the values from the two distributions.

To further check the adequacy of the fitted probability distribution, the autocorrelation of the errors of both the real and simulated probability distributions were analyzed. Figs. 10 and 11 shows the autocorrelation of the errors for the real and simulated probability distributions respectively.

The Box-Ljung test shows that the autocorrelations among the errors for the real probability distribution are zero (*p*-value = 0.30). This indicates that the errors are random and that the probability distribution provides an adequate fit to the data. Likewise, the Box-Ljung test shows that the autocorrelations among the errors for the simulated probability distribution are zero (*p*-value = 0.43). This indicates that the errors are random and that the log-normal probability distribution provides an adequate fit

of 0.91 and log standard deviation of 0.46. The simulated Lognormal data was then compared with the original data (the exchange rate between the Ghana Cedi and the American dollar). Fig. 7 show plots of the original data on the exchange rate between the Ghana Cedi and the American dollar. Fig. 8 also show the plots from the simulated lognormal data. From Figs. 7 and 8, it can be observed that the two datasets have similar properties. Fig. 9 also shows the plot from these two datasets. From these graphs in Figs. 7–9, it can be observed that the two datasets are comparable with each other. From the empirical and theoretical cumulative distribution function the Q-Q and P-P plots, almost all the points lie along the fitted curve. Hence, the Lognormal distribution is the best fit for the data.

The Wilcoxon signed rank test with continuity correction was conducted to test if the two datasets have the same distribution
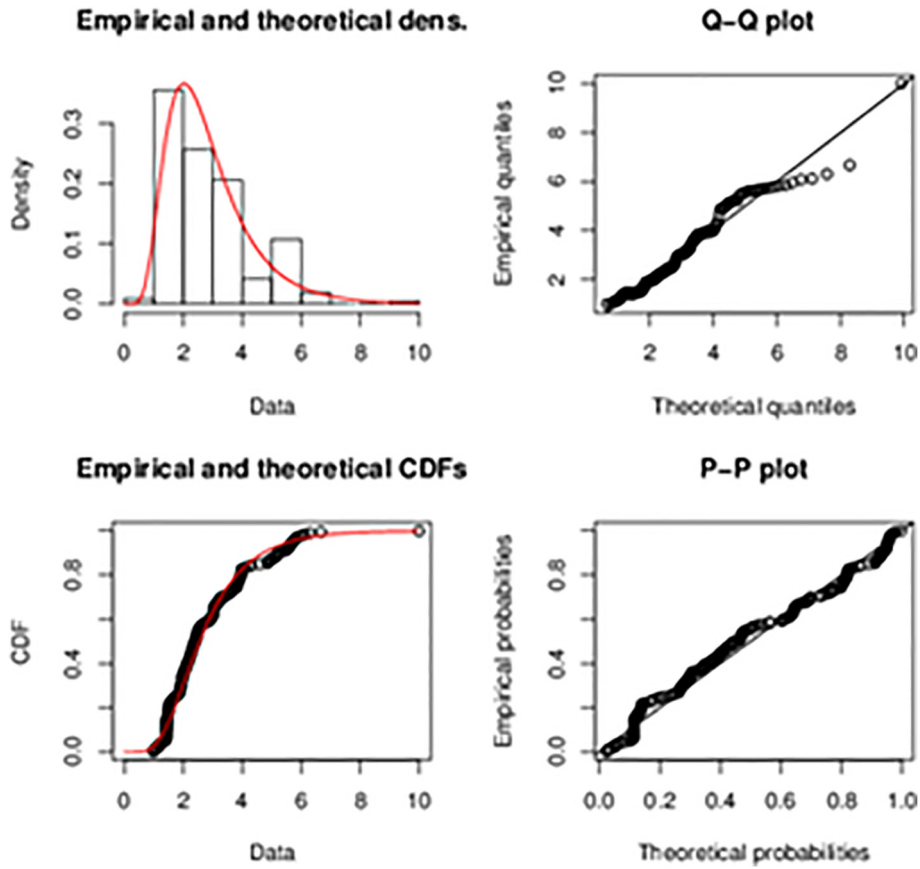
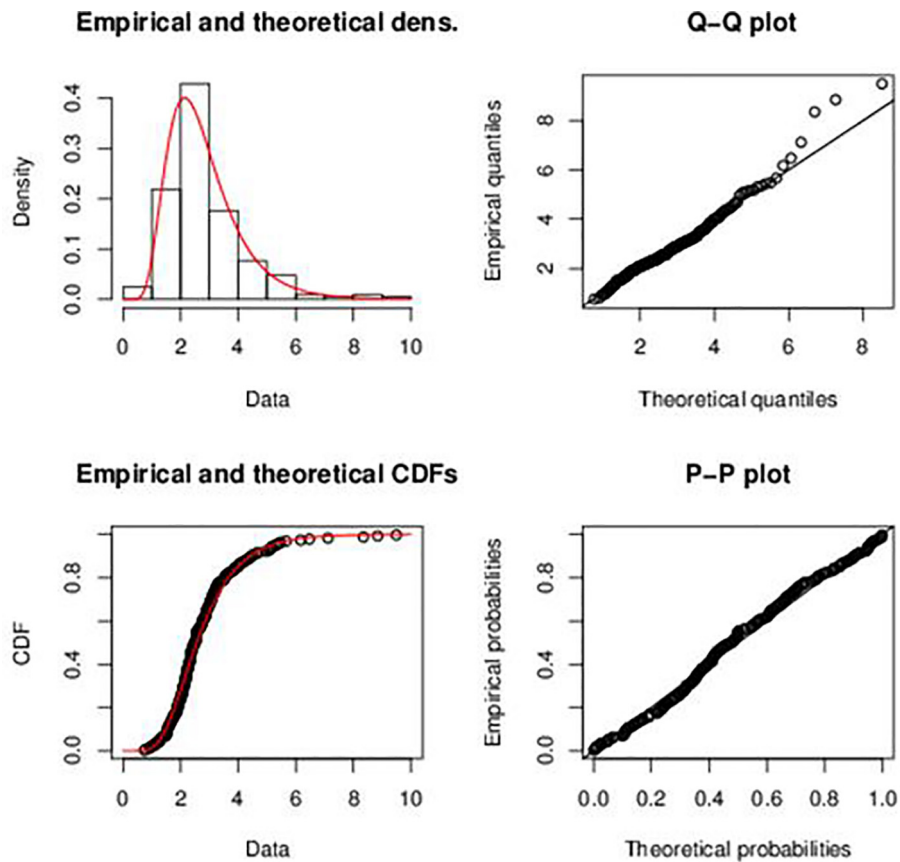**Fig. 7.** Plots of the original data on the exchange rate between the Ghana Cedi and the American dollar.

Fig. 8. Plots of the simulated data of the exchange rate between the Ghana Cedi and the American dollar.
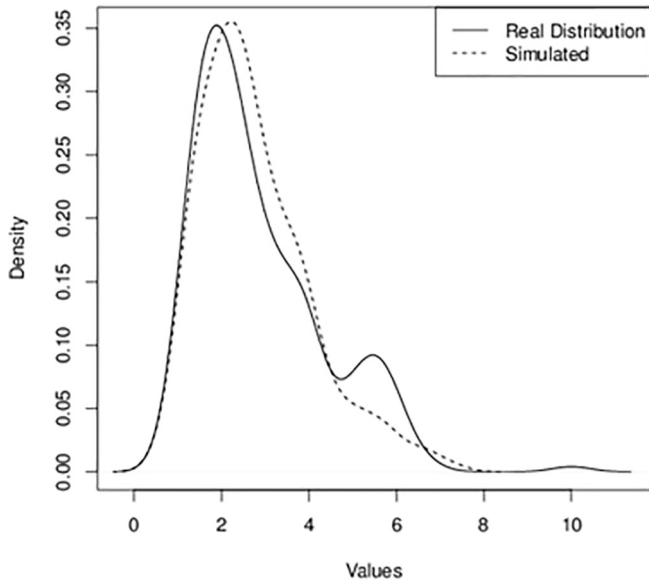
Fig. 9. A graph of the distribution of real and simulated exchange rate between the Ghana Cedi and the American dollar.

to the data. Furthermore, the Kullback-Leibler divergence was calculated to assess the amount of information lost when any of the probability distributions was used to approximate the data. Table 2 show the probability distributions and the corresponding Kullback-Leibler divergence values calculated. It can be observed

that, less amount of information is lost when the lognormal probability distribution is used to approximate the data.

3.1. Gaussian mixture models

Table 3 shows the summary statistics of the three possible clusters in the data. It can be observed that, when the data is considered to be a single cluster, the log-likelihood, Bayesian Information Criterion and Integrated Complete-data Likelihood had the highest values. This implies that, in spite of the fact that the data maybe having some clusters inherent in it, the data is well represented when it is assumed that the data has only one cluster.

4. Conclusion

The sole aim of persons who deal in exchange of currencies is to minimize loss through the volatilities in the exchange rates between the currencies. Hence, being able to forecast accurately the exchange rate maximizes the returns in profit. Hence, this research has adequately proved that the exchange rate between the Ghana Cedi and the American dollar follows the lognormal distribution. By using the Lognormal distribution in predicting the possible exchange rate between the two currencies will result in maximum returns in profit. From the autocorrelation of the errors of both the real and simulated probability distributions in addition to the Box-Ljung test, it was observed that the errors are random. The randomness of the autocorrelation of the errors shows that the fitted distribution adequately fits the data. From the Kullback-Leibler divergence values calculated, it can be observed that, there is less loss of information when the lognormal probability
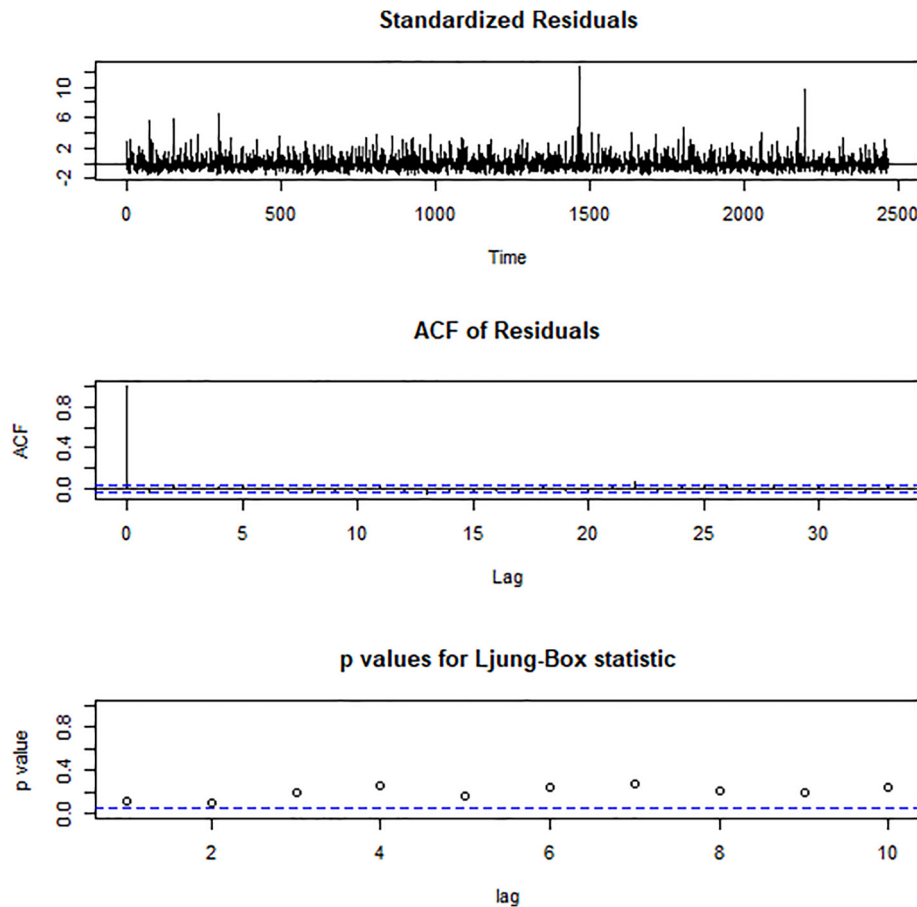

Fig. 10. The autocorrelation of the errors for the real probability distribution.
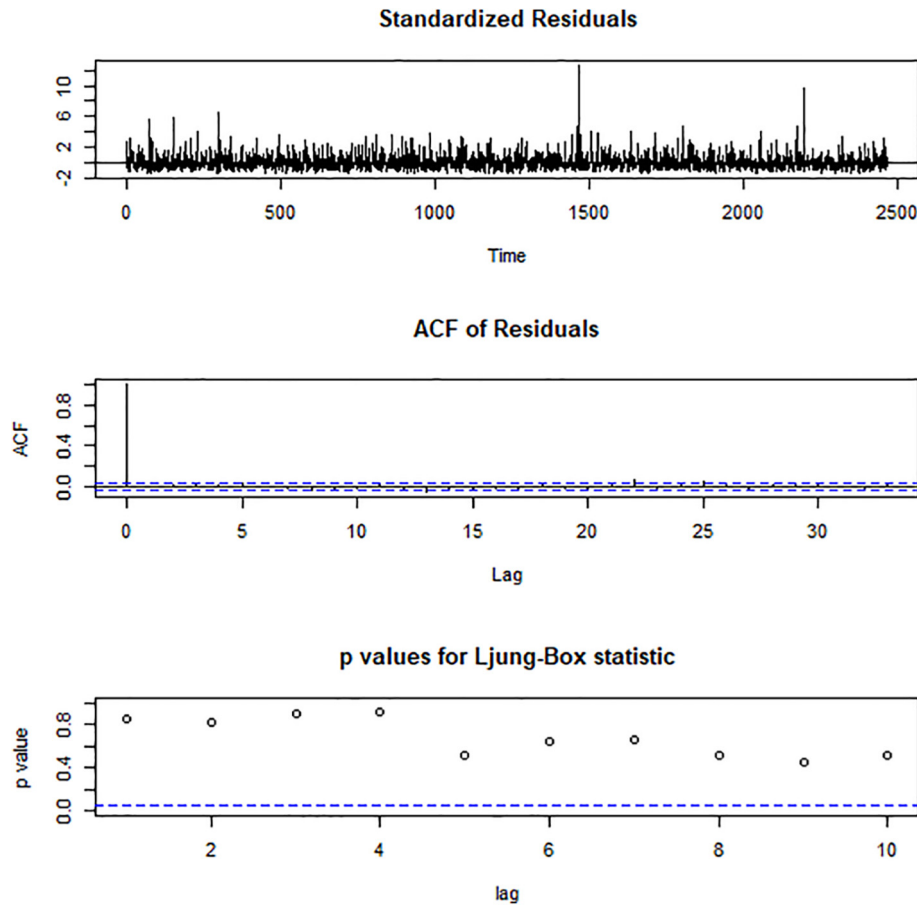
Standardized Residuals

ACF of Residuals

p values for Ljung-Box statistic

Fig. 11. The autocorrelation of the errors for the simulated probability distribution.

Table 2
The Kullback-Leibler divergence values for the exchange rate between the Ghana Cedi and the American dollar.

Probability distribution	Kullback-Leibler divergence
Weibull	0.65
Gamma	0.79
Lognormal	0.26
Exponential	0.63

Table 3
Summary statistics for Gaussian mixture model.

Number of clusters	Log-likelihood	Bayesian information criterion	Integrated complete-data likelihood
1	−5186.83	−10389.28	−10389.3
2	−5038.97	−10116.99	−10347.7
3	−4929.5	−9921.51	−10327.8

distribution is used to approximate the data than the other probability distributions. Hence, the lognormal probability distribution better approximates the data.

References

Akgiray, V., Booth, G.G., 1987. Compound distribution models of stock returns: an empirical comparison. J. Fin. Res. 10 (3), 269–280.

Amado P., 2002. Skewness in Individual Stocks at Different Investment Horizons, 2 Quantitative Finance 139.

Andersen, T.G., Diebold, F.X., Bollerslev, T., Labys, P., 2001. The distribution of realized exchange rate volatility. American statistical association. J. Am. Stat. Assoc. 96 (453). Applications and Case Studies.

Aparicio F. M. and Estrada J., 1996. Empirical distributions of stock returns: Scandinavian securities markets, 1990-95. Departamento de Estadística Working Papers 1996-10. Statistics and Econometrics. Universidad Carlos III de Madrid.

Aron, J., Elbadawi, I., Kahn, B., 1998. Determinants of the real exchange rate in South Africa, CSAE Working Paper Series 1997-16. Centre for the Study of African Economics. University of Oxford.

Bai, X.-Z., Russell, J.R., George, C., Tiao, G.C., 2003. Kurtosis of GARCH and stochastic volatility models with non-normal innovations. J. Econ. 114 (2), 349–360.

Brooks, C., 2002. Introductory Econometrics for Finance. Cambridge University Press.

Corrado, C.J., Zivney, T.L., 1992. The specification and power of the sign test in event study hypothesis tests using daily stock returns. J. Fin. Quantit. analysis 27 (3), 465–478. Cambridge University Press.

Carroll, R.J., 2014. Estimating the distribution of dietary consumption patterns. Stat Sci. 29 (1), 2–8.

Chang, P.H.K., Melick, W.R., 1999. Workshop on estimating and interpreting probability density functions. Bank for Intrnational Settlements Information. Press Library Serv., 1–10

Ding, Y., Li, S., Li, L., 2009. An analysis on chaos behavior of currency exchange rate undulation. First International Workshop on Education Technology and Computer Science, Wuhan, Hubei, pp. 599–602.

Egan W. J., 2007. The Distribution of S&P 500 Index Returns.

Fama, E.F., 1976. Foundations of finance. J. Fin. 38, 143–145.

Fujihara, R., Park, K., 1990. The probability distribution of future price in the foreign exchange market: a comparison of candidate processes. J. Future Markets 10 (6), 623–641.

Gradojevic, N., Yang, J., 2006. Non-linear, non-parametric, non-fundamental exchange rate forecasting. J. Forecast. No. 25, 227–245.

Huang, S.C., Chuang, P.J., Wu, C.F., 2010. Chaos-based support vector regressions for exchange rate forecasting. Expert Syst. Appl. 37 (12), 8590–8598.

Lin, C.-S., Chiu, S.-H., Lin, T.-Y., 2012. Empirical mode decomposition-based least squares support vector regression for foreign exchange rate forecasting. Econ. Model. 29, 2583–2590.

Radhwan, A., Kamel, M., Dahab, M.Y., Hassanien, A., 2015. Forecasting exchange rates: a chaos-based regression approach. Int. J. Rough Sets Data Analysis 2 (1), 38–57.

Tucker, A.L., Pond, L., 1988. The probability distribution of foreign exchange price changes: test of candidate processes. Rev. Econ. Stat. 70 (4), 638–647.