

# EMPIRICAL DISTRIBUTIONS OF STOCK RETURNS - GHANA STOCK EXCHANGE

Stephen Opoku Oppong  
Department of IT  
Academic City College, Accra  
sopokuoppong@yahoo.com

Dominic Asamoah  
Department of Comp. Sci.  
KNUST, Kumasi  
dominic\_asamoah@yahoo.co.uk

Emmanuel Ofori  
Oppong  
Department of Comp. Sci.  
KNUST, Kumasi  
eoforioppong@yahoo.com

## ABSTRACT

The normality assumption about the distribution of stock returns has long been challenged both empirically and theoretically. Alternative distributions have been proposed to better capture the characteristics of stock return data. Finding the appropriate probability distribution is relevant especially in risk management techniques which rely heavily on the probability distribution of data. This article investigates whether the Gaussian(Normal) distribution hypothesis holds for stock returns traded on the Ghana Stock Exchange and also the ability of three alternative distributions (Logistic distribution, Three parameter Logistic distribution and Three parameter Lognormal distribution) to represent the behavior of daily stock returns over the period from 2007 to 2015. The Maximum Likelihood Estimate and Method of Moments is used to estimate parameters of the proposed probability distributions and the Kolmogorov-Smirnov test and the Anderson-Darling test is used to perform Goodness of Fit test on the parameters estimated. The normality assumption is out rightly rejected by the goodness of fit test used and the Three parameter Loglogistic distribution and Logistic distribution fits the daily stock returns best.

*Keywords: probability distribution, normality, goodness of fit, returns*

## I. INTRODUCTION

The assumption that stock returns are normally distributed is widely used, implicitly or explicitly, in theoretical finance (asset pricing and option pricing models, random walk theory etc.). The shape of the tail of the returns distribution is having more implications rather than ignored by some economists in the past.

Various authorities have successfully fitted assets returns to probability distribution. Alfonso et al. [1] found out that, despite fact that the normal distribution simplifies

analytical calculation which is very valuable, empirical studies by Mandelbrot [13],

Mantegna and Stanley [14], Gopikrishnan et al. [8] and Cont [6], showed that the distribution of returns of stock prices have a fatter tail than that of a Normal distribution. A detailed statistical analysis of logarithm returns of daily fluctuations of the IPC; the leading Mexican Stock Market Index, shows that the stable Levy distribution best fits the data better than the Normal, Normal Inverse Gaussian and Levy distributions [1]. Aparicio and Estrada [4] found that the scaled-t distribution had an overall support where ten of the fourteen markets of the European stock market analyzed was best fitted by it but the Normal distribution displayed the worst fit and neither the logistic nor the exponential power distributions gave a good fit to the empirical distributions. Research has found that financial returns experience fat tails, which suggests that the normal distribution functions well in predicting foreseeing regular but is not a good estimator to predict extreme events [7]. Blattberg and Gonedes [5] suggested using the Student (or t) distribution to account for the fat tails of return distributions observed in earlier studies. Olson et al [14] essential result of their research affirms broadly expressed opinions that financial return data has fat tails, thus, not normally distributed and is better fitted by the logistic distribution. The effect is that the risk is understated when the normality assumption is used. The logistic distribution fitted best in most cases matching expectations that it includes the ability to fatten tails. The student-t distribution had the second best fit in most cases, while normal and lognormal were either third or fourth in all cases. Daniel Lee Richey [16] also confirmed that returns for individual companies and even indices, do not follow a normal distribution. The Student-t distribution was the best but when analyzing

data with smaller intervals and the 3-parameter lognormal distribution was the best, especially for the 1-month intervals. Meelis Käärrik and Anastassia Žegulova [10] in a follow-up to their preliminary research [9] where they established that lognormal distribution had best fit among the candidates realized that the lognormal assumption was too strong and the tail behaviour needs to be revisited. Using data provided by Estonian Traffic Insurance Fund focused on a model where the main part of the data follows a (truncated) lognormal distribution and the tail is fitted by generalized Pareto distribution.

This paper seeks to find out the most suitable distribution for the stocks returns on the Ghana Stock Exchange first by testing the normality assumption for daily stock returns and attempting to find the specification that best fits the data in each market.

**II. METHODOLOGY**

The distributions under consideration are 3-parameter Lognormal, Logistic, 3-parameter Loglogistic and Gaussian (Normal) distributions.

Techniques for estimating the parameters of the distributions which includes the Maximum Likelihood Estimator and the Method of Moments would be discussed. The Kolmogorov-Smirnov and Anderson-Darling goodness of fit tests applied in this research work are also discussed.

**Probability Distribution**

The candidate distributions (3-parameter Lognormal, Logistic, 3-parameter Loglogistic and Gaussian (Normal)) are discussed below.

**Normal Distribution**

It is the most popular distribution in various financial, physical and engineering applications. The probability density function of the normal distribution is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

(1)

for  $-\infty < x < \infty$ , where  $-\infty < \mu < \infty$  and  $0 < \sigma < \infty$

**Three-Parameter Lognormal Distribution**

The probability density function (pdf) of the three-parameter lognormal distribution is

$$f(x; \mu, \sigma, \gamma) = \frac{1}{(x-\gamma)\sigma\sqrt{2\pi}} \exp\left\{-\frac{[\ln(x-\gamma)-\mu]^2}{2\sigma^2}\right\}$$

(2)

where  $x > \gamma \geq 0$ ,  $-\infty < \mu < \infty$  and  $\gamma$  is the threshold parameter or location parameter that defines the point where the support set of the distribution begins,  $\mu$  is the scale parameter that stretch or shrink the distribution and  $\sigma$  is the shape parameter that affects the shape of the distribution.

**Logistic Distribution**

The probability density function (pdf) of the logistic distribution is given by:

$$f(x / \mu, \sigma) = \frac{e^{-\left(\frac{x-\mu}{\sigma}\right)}}{\sigma \left(1 + e^{-\left(\frac{x-\mu}{\sigma}\right)}\right)^2}$$

(3)

where  $-\infty < \mu < \infty$ ,  $\sigma \geq 0$  and  $-\infty < x < \infty$ .

Logistic distribution is a continuous probability distribution.

**Three-Parameter LogLogistic Distribution**

The probability density function (pdf) of the three-parameter loglogistic distribution is

$$f(x) = \frac{\alpha}{\beta} \left(\frac{x-\gamma}{\beta}\right)^{\alpha-1} \left(1 + \left(\frac{x-\gamma}{\beta}\right)^\alpha\right)^{-2}$$

(4)

where  $\alpha, \beta > 0$ ,  $\gamma \leq x < \infty$ .  $\alpha$  is the shape parameter,  $\beta$  is the scale parameter and  $\gamma$  the location parameter.

**Methods of Point Estimation**

The two methods used for estimation of the parameters of the candidate probability distribution are Maximum Likelihood Estimation and Method of Moments.

**Maximum Likelihood Estimation (MLE)**

**Method**

Let X be a continuous random variable with pdf (probability density function  $pdf = f(x; \theta_1, \theta_2, \dots, \theta_k)$  where  $\theta_k$  are the unknown parameters for the N independent observations  $x_1, x_2, \dots, x_N$ .

The Likelihood Function is

$$L(x_1, x_2, \dots, x_N / \theta_1, \theta_2, \dots, \theta_k) = \prod_{i=1}^N f(x_i; \theta_k) \tag{5a}$$

and the logarithmic likelihood function  $\Lambda$  is;

$$\Lambda = \ln L(x_1, x_2, \dots, x_N / \theta_1, \theta_2, \dots, \theta_k) \tag{5b}$$

$$= \sum_{i=1}^N \ln f(x_i; \theta_1, \theta_2, \dots, \theta_k) \tag{5c}$$

MLE of  $\theta_1, \theta_2, \dots, \theta_k$  are obtained by maximizing L or  $\Lambda$

**Method of Moments**

Karl Pearson introduced the method of moments in 1894. The method of moments is a method of estimation of population parameters. One starts with deriving equations that relate the population moments to the parameters of interest. The equations are then solved for the parameters of interest, using the sample moments in place of the (unknown) population moments. This results in estimates of those parameters. The steps are defined below

- Let  $x_1, x_2, \dots, x_n$  be a sample random selected from the probability distribution  $f(x, \theta)$ , where  $\theta = (\theta_1, \theta_2, \dots, \theta_k) \in \Omega, k \geq 1$

is/are parameter(s) and  $\Omega$  is the sample size of the parameters

- Obtain the first k moments about the origin,  $\mu_j(\theta)$  and its corresponding

sample moments,  $m_j$ , defined by

$$u_j(\theta) = E(x^j) = \int_{R_x} x^j f(x, \theta) dx \tag{6a}$$

$$m_j = \frac{1}{n} \sum_{i=1}^n x_i^j \text{ where } j = 1, 2, \dots, k \tag{6b}$$

- Set up the equation(s),  $u_j = m_j$ , where  $j = 1, 2, \dots, k$  and obtain the unique solution(s),  $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k$ .

**Goodness of Fit Tests**

The Goodness of fits test which is based on statistical theory is used to determine whether

data are consistent with a given probability distribution, i.e. to test the hypothesis that an observed probability distribution fits (or conforms to) some claimed distribution,

Two(2) Goodness of fits test are discussed namely, Kolmogorov-Smirnov (KS), Anderson Darling (AD).

**Kolmogorov-Smirnov Goodness of Fit Test**

The Kolmogorov-Smirnov (KS) test, introduced by Kolmogorov and Smirnov [11], [12] is a test of the distance or deviation of empirical distributions from a required theoretical distribution

The KS-statistic is sometimes abbreviated as D-statistic.

The KS statistic for a given theoretical cumulative distribution  $F(x)$  is

$$D = \sup_x |F_n(x) - F(x)| \tag{7}$$

where:

- $F(x)$ : theoretical cumulative distribution value at x
- $F_n(x)$ : empirical cumulative distribution value for a sample size of n.

The null hypothesis that  $F_n(x)$  comes from the underlying distribution  $F(x)$  is rejected if D-statistic is larger than the critical value  $D_\alpha$  at a given  $\alpha$

**Anderson-Darling Test**

The Anderson-Darling test, developed by T.W. Anderson and D.A. Darling [2], [3] in 1952 was an alternative to other statistical tests for detecting sample distributions' deviation from normality.

The one-sample AD test statistic is non-directional, and is calculated from the following

formula:

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) \left( \ln F(X_{(i)}) + \ln(1-F(X_{n+1-i})) \right) \tag{8}$$

where

F is the cumulative distribution function (cdf) of the specified distribution

$X_i$  is the ordered data.

The hypothesis regarding the distributional form is rejected at the chosen significance

level ( $\alpha$ ) if the test statistic  $A^2$ , is greater than the critical value  $AD_\alpha$  at a given  $\alpha$

**Data and Test For Normality**

The sample under consideration consists of thirty-eight (38) stocks traded on the Ghana Stock Exchange namely AngloGold Ashanti Limited(AADS), African Champion Industries Limited(ACI), AngloGold Ashanti Limited(AGA), Aluworks Ltd (ALW), Ayrton Drugs Manufacturing Company LTD (AYRTN), Benso Oil Palm Plantation Limited (BOPP), CAL Bank Limited(CAL), Clydestone (Ghana) Limited(CLYD), Camelot Ghana Ltd.(CMLT), Cocoa Processing Company(CPC), Ecobank Ghana Limited(EBG), Enterprise Group Limited(EGE), Ecobank Transnational Incorporation(ETI), Fan Milk Ltd(FML), Ghana Commercial Bank Ltd.(GCB), Guinness Ghana Breweries Ltd.(GGBL), NewGold issuer Ltd(GLD), Ghana Oil Company Limited(GOIL), Golden Star Resources Ltd(GSR), Golden Web Ltd.(GWEB), HFC Bank Ltd(HFC), Mega African Capital(MAC,Mechanical Lloyd Company Ltd(MLC), Produce Buying Company Ltd.(PBC), Pioneer Kitchenware Ltd.(PKL), PZ Cussons (Ghana) Ltd(PZC), Standard Chartered Bank (Ghana) Ltd.(SCB), Standard Chartered Bank (Ghana) Ltd Preference Shares (SCB PREF), SIC Insurance Company Limited(SIC), Societe Generale Ghana(SOGEH), Starwin Products Limited(SWL), Sam Wood Ltd.(SWL), Trust Bank Ltd(TBL), Tullow Oil(TLW), Total Petroleum Ghana Ltd(TOTAL), Transol Solutions (Ghana) Limited(TRANSOL), Unilever (Ghana) Ltd.(UNIL), UT Bank(UTB)

The period under consideration starts from June 25, 2007, to December 31, 2015. The series of returns is defined as

$$R_t = \frac{N_t - N_{t-1}}{N_{t-1}}$$

(9)

where

$R_t$ , the return  
 $N_t$  the Stock Price in day  $t$ .

Table 1 summarizes some relevant information about the empirical distributions of stock returns under consideration.

The statistics reported are the mean, standard deviation, minimum and maximum return during the sample period, skewness and kurtosis and sample size

Generally, for a distribution to be perfectly normally distributed, values for skewness and kurtosis must be zero (0) and three (3) respectively. The last column (column 8) of Table 3.1 shows that all of the financial stocks under consideration are clearly leptokurtic, thus exhibiting high peaks.

Table 2 shows the results of Goodness of Fit test for AD and KS and their p-value (threshold value).

The p-value used to test for normality all rejected the null hypothesis that the distributions for the financial stocks are normally distributed at all significance levels tested from 0.1 to 0.001, since the significance level is higher than the p-value.

Even though these tests use different information, the results of all of them directs us to the same point i.e. to the rejection of the assumption of normality for the daily stock returns

**Parameters Estimated And Goodness Of Fit Test**

Table 3 shows parameters estimated using Maximum Likelihood Estimate and Method of Moments

In order to compare the relative fit of the theoretical distributions considered, AD and KS goodness-of-fit tests are employed

The results of these tests are shown below in Table 4 and Table 5.

**V. DISCUSSION**

The summary of the Goodness of Fit Test is shown in Table 6. It can be seen that the Loglogistic(3P) distribution fit 14 out of the 38 stocks returns considered best whiles the AD test fitted the Logistic distribution best to 28 stock returns.

**IV. CONCLUSION AND FUTURE WORKS**

It is concluded that the Three Parameter Loglogistic distribution fits the returns of the stocks traded on the Ghana Stock Exchange using the KS test and the Logistic distribution using the AD test better than the Normal distribution which is predominantly used in financial markets. This is because the AD test gives more emphasis to the tail of the distribution.

It is recommended that future studies should be conducted on the tail decay of the return distributions of Ghanaian financial data and also mixture of probability distribution to fit stock returns

## V. REFERENCES

- [1] Alfonso, L., Mansilla, R. and Terrero-Escalante, A.C.(2011). *On the scaling of the distribution of daily price fluctuations in Mexican financial market index*. Physica A.
- [2] Anderson, T. W., Darling, D. A.(1952) *Asymptotic theory of certain goodness-of-fit criteria based on stochastic processes*. Annals of Mathematical Statistics, 23, 193-212.
- [3] Anderson, T. W., Darling, D. A.(1954) *A test of goodness of fit*. Journal of the American Statistical Association, 49, 765–769,
- [4] Aparicio, Felipe, and Javier Estrada. (2001) *Empirical Distributions of Stock Returns: European Securities Markets, 1990–95*. European Journal of Finance, 7, pp. 1–21.
- [5] Blattberg, R. and N. Gonedes (1974). *A Comparison of the Stable and Student Distributions as Statistical Models for Stock Prices*. Journal of Business, 47, 244-280.
- [6] Cont, R. (2001). *Empirical properties of asset returns: stylized facts and statistical issues*. *Quantitative Finance*, 1, pp: 223-236.
- [7] Dowd, K. (1998), *Beyond Value At Risk: The new science of risk management*. West Sussex: John Wiley & Sons Ltd.
- [8] Gopikrishnan, P., Meyer, M., Amaral, L. A. N. and Stanley, H. E. (1999). *Inverse cubic law for the distribution of stock price variations*. European Physics Journal B, 3(2), 139–140.
- [9] Käärik, M., and Umbleja, M. (2010), *Estimation of claim size distributions in Estonian traffic insurance; In: Selected Topics in Applied Computing*. Proceedings of Applied Computing Conference (ACC '10), Timisoara, Romania, pp. 28–32.
- [10] Käärik M. and Žegulova A, (2012) *On Estimation Of Loss Distributions And Risk Measures*, Acta Et Commentationes Universitatis Tartuensis De Mathematicavolume 16, Number 1
- [11] Kolmogorov, A. N. (1933), *Sulla Determinazione Empirica Di Une Legge Di Distribuzione*. *Giornale Dell'istituto Italiano degli Attuari*, 4, 83-91,
- [12] Kolmogorov, A. N. *Confidence limits for an unknown distribution function*. Annals of Mathematical Statistics, 12, 461-463, 1941
- [13] Mandelbrot, B. (1963). *The variation of certain speculative prices*. Journal of Business, 36: 394–419.
- [14] Mantegna, R. N. and Stanley, H. E. (1995). *Stochastic processes with ultraslow convergence to a Gaussian: The truncated Levy flight*. Physical Review Letters, 73, 2946–2949.
- [15] Olson, David L and Desheng Wub *The impact of distribution on value-at-risk measures*, Mathematical and Computer Modelling 58 (2013) 1670–1676
- [16] Richey, D.L (2012), *The Distribution Of Individual Stock Returns In A Modified Black-Scholes Option Pricing Model*.

## AUTHORS PROFILE



**Mr. Stephen Opong** is a Lecturer in the Department of Information Technology, Faculty of Technology at Academic City College, Accra Ghana. He attended Kwame Nkrumah University of Science and Technology (KNUST), Kumasi where he read Statistics and Actuarial Science for his BSc degree. He pursued further studies and was awarded a Master of Philosophy (M.Phil) in Information

Technology also by KNUST. He has undertaken researches in risk management, statistical modeling and currently engaged in research in computer graphics and image processing.



**Mr. Dominic Asamoah** is a Lecturer in the Department of Computer Science, KNUST. He possesses a BSc and MPhil Degree in Computer Science from KNUST. Research areas include Image processing and Data Structures.



**Mr. Emmanuel Ofori Oppong** is a Lecturer in the Department of Computer Science, KNUST. He possesses a BSc degree in Computer Science and MSc in Industrial Mathematics from KNUST. His research areas include optimization and programming.