Vol 4, No.8 August 2014 ISSN 2225-7217
ARPN Journal of Science and Technology
©2011-2014. All rights reserved.

Three Parameter Log normal Multidimensional Diffusion Process With Exogenous Factors

Manal Bouskraoui[1], Aziz Arbai[2]
[1]University of Abdelmalek Essaadi, Morocco email: bouskraoui@gmail.com
[2]Department of mathematics and computer,
University of Abdelmalek Essaadi, Morocco email: arbai_aziz@yahoo.com

***Abstract*—blackThis article we propose a log normal distribution with a threshold parameter. This was considered as an extension of the bi-parameter process by adding third parameter. The Kolmogorov equations and Ito's stochastic differential equations give us the transition probability density function and the moments of this process. Assuming the discrete sampling of the sample path of the model, the local maximum-likelihood estimate the parameters except the threshold, which it estimation requires us to solve a non-linear equation for this, Newton-Raphson method is called.**

***Index Terms*—blackThree-parameter; log normal process; discrete sampling; diffusion process; local maximum-likelihood estimation.**

I. Introduction

Log normal distribution with three-parameter is one of the distributions commonly used in several field, biology, geology, agriculture, statistics, Economy and operations Research. The most celebrated applications are turn about studying Concentration of antibodies in the blood by Royston[21], modeling distribution volume of a raindrop, Queuing theory, recently, the study of the acidification lakes by Crawford[8] and Tintner[23] who has devoted the research on the volatility of the stock market. The log normal distribution was invented in 1879 by McAlister, the first mathematician was setting the main pillars of the theory. Since then, his theory has become a controversial topic attracting scientists. In 1917 Wicksell devoted these studies to Age distribution of the first marriages. These observations led to introduce a new parameter, which take threshold distribution name. The reason for the log normal distribution had an extension by the introduction of third parameter. The theoretical difficulty which we face by applying the three-parameter log normal distribution is the use the parameters estimations methods; method of maximum likelihood, iterative methods...Indeed, as soon as the likelihood function tends to the global maximum (+ infinity), the parameters tend to inadmissible values. Hill's article[15] at 1963, we exhibited a Bayesian argument to justify the use estimator parametric where the likelihood function tend to largest local maximum. His observations were developed by Griffiths[12], he confirmed in non-Bayesian viewer, that the local likelihood estimator is a good estimator. Since then, a lot of study and results have been published on using the numerical methods to calculate local likelihood estimator and estimate log normal distribution parameters, the idea is to find a stationary point of the likelihood function , by using the Newton Method or other iterative methods, The article of Cohen [5] Lambert[16], Harter and Moore[13]and Calitz [4] explain more details. There are two different approaches to considerate of diffusion process: the first takes the diffusion process as being solution of an Ito stochastic differential equation and estimates the parameters by using continuous sampling by the maximum-likelihood method as described in Brown and Hewitt[3] Basawa and Prakasa Rao [2], Molina, Hermoso [18] and Gutiez, Hermoso and Molina [10]. Whilst the second approach considers the diffusion process as being the solution of kolomogorov equations, it estimates the parameters by using discrete sampling with the maximum-likelihood method, as described Tintner [24][25], Gutiérrez, Angulo, González and Pérez [11], Torres[22] and Arbai[1].

II. The three-parameter log normal process with q exogenous factors:

A. Definition

We introduce the three-parameter log normal multidimensional diffusion process with q exogenous factors, using the forward and backward Kolomogorov equations.
Let $\gamma = (\gamma_1, ..., \gamma_k)$ a k-dimensional vector and consider $X(t)$ a Markov almost surly continuous $\prod_{i=1}^{k}]\gamma_i, +\infty[$ valued sampling process with parametric space $[t_0, T]$ with transition probability given by:

$$P(y, t/x, s) = P[X(t) = y/X(s) = x]$$

where $X(t) = (X_1(t), ..., X_k(t))\prime$ $X(s) = (X_1(s), ..., X_k(s))\prime$; x and y are the k-dimensional vectors of associated variables.
We assume the following conditions:

1) $\lim_{h \to 0} \frac{1}{h} \int_{|y-x|>\epsilon} P(d_y, t+h/x, t) = 0$
2) $\lim_{h \to 0} \frac{1}{h} \int_{|y-x|\leq\epsilon} (y-x) P(d_y, t+h/x, t) = f(x, t) =$

$$\begin{pmatrix} (f_{10} + \sum_{j=1}^{q} f_{1j} F_j(t)(x_1 - \gamma_1) \\ \vdots \\ \left(f_{k0} + \sum_{j=1}^{q} f_{kj} F_j(t)\right)(x_k - \gamma_k) \end{pmatrix}$$

where $F_j, j = 1, ..., q$ are continuous functions over $[t_0, T]$.

©2011-2014. All rights reserved.

3) $\lim_{h\to 0}\frac{1}{h}\int_{|y-x|\leq\epsilon}(y-x)(y-x)\prime P(d_y,t+h/x,t)=(x-\gamma)B(x-\gamma)\prime$
where B is definite, non negative, symmetric matrix with elements $b_{ij}>0$,
4) The infinitesimal moments similar to the former of order three or more vanishing.

B. Definition

Let $F_1(t),F_2(t),...,F_q(t)$ a exogenous factors, $Z(t)=X(t)-\gamma$ the exogenous factors affect only the drift but not the covariance matrix, Under these conditions, the backward and the forward kolomogorov equations can be obtained through the diffusion characterized by 2 and 3.

$$\frac{\partial P}{\partial s}+\frac{1}{2}\sum_{k}^{i,j=1}b_{ij}(x_i-\gamma_i)(x_j-\gamma_j)\frac{\partial P}{\partial x_i\partial x_j}+\sum_{i=1}^{k}\left(f_{i0}+\sum_{q}^{l=1}f_{il}F_l(t)\right)$$

$$(x_i-\gamma_i)\frac{\partial P}{\partial x_i}=0$$

$$\frac{\partial P}{\partial t}+\frac{1}{2}\sum_{k}^{i,j=1}b_{ij}\frac{\partial^{2[(y_i-\gamma_i)(y_i-\gamma_j)P]}}{\partial_{y_i}\partial_{y_j}}-\sum_{i=1}^{k}\left(f_{i0}+\sum_{i=1}^{q}f_{il}F_l(t)\right)$$

$$\frac{\partial[(y_i-\gamma_i P]}{\partial y_i}=0$$

where $x=(x_1,...,x_k)\prime$ and $y=(y_1,...,y_k)\prime$ are such $x_i>\gamma_i$ and $y_i>\gamma_i$ for $i=1,...,k$,
take $P=P(y,\frac{t}{x},s)$ a transition conditional density,
the common solution to these equations with initial condition

$$P=P(y,s/x,s)=\delta(y-x)$$

take the form

$$P(y,t/x,s)=\left[\left(\prod_{i=1}^{k}(y_i-\gamma_i)\right)(2\Pi)^{\frac{k}{2}}(t-s)^{\frac{k}{2}}|B|^{\frac{1}{2}}\right]^{-1}$$

$$\exp\left\{-\frac{Q}{2(t-s)}\right\}$$

where Q is the following quadratic form:

$$Q=\left(\ln(y-\gamma)-\ln(x-\gamma)-\beta_0(t-s)-\sum_{i=1}^{k}\beta_i\int_s^t F_i(\tau)d\tau\right)\prime B^{-1}$$

$$\times\left(\ln(y-\gamma)-\ln(x-\gamma)-\beta_0(t-s)-\sum_{i=1}^{k}\beta_i\int_s^t F_i(\tau)d\tau\right)$$

with

$$\beta_0=(f_{10}-\frac{1}{2}b_{11},...,f_{k0}-\frac{1}{2}b_{kk})\prime=f_0-\frac{1}{2}diag(B)$$

and

$$\beta_i=(f_{1i},...,f_{ki})\prime,i=1,...,q.$$

C. Moments of the process

The moments of the three-parameter log normal multidimensional diffusion process are given by the moments of the bi-parameter multidimensional diffusion process.
Let $Z(t)=W(t)-\gamma$, so if

$$X(s)=x_s=(x_{1s},...,x_{ks})\prime$$

$$E[\exp^{\sigma'\ln(z(t))}]=\exp\left\{\sigma'\left(\ln(x_s-\gamma)+\beta_0(t-s)+\right.\right.$$

$$\sum_{i=1}^{q}\beta_i\int_s^t F_i(\tau)d\tau+\frac{1}{2}(t-s)\sigma' B\sigma$$

With,

$$\sigma'=(\sigma_1,....,\sigma_k)$$

And,

$$E[X_i(t)]=\gamma_i+(x_{is}-\gamma_i)\exp\left(f_{i0}(t-s)+\sum_{j=1}^{q}\beta_{ij}\right.$$

$$\int_s^t F_j(\tau)d\tau when\, i=1,...,k$$

$$Cov[X_i(t),X_j(t)]=(x_{is}-\gamma_i)(x_{js}-\gamma_j)\exp\left(f_{i0}(t-s)+\right.$$

$$\sum_{j=1}^{q}\beta_{il}\int_s^t F_l(\tau)d\tau\times\exp\left(f_{i0}(t-s)+\right.$$

$$\sum_{j=1}^{q}\beta_{jl}\int_s^t F_l(\tau)d\tau$$

$$[\exp(b_{ij}(t-s))-1]$$

$$Var[X_i(t)]=(x_{is}-\gamma_i)^2\exp\left(2f_{i0}(t-s)+2\sum_{j=1}^{q}\beta_{il}\right.$$

$$\int_s^t F_l(\tau)d\tau$$

$$\times\exp[(b_{ij}(t-s))-1]$$

III. Estimation of the parameters

We estimate the parameters of the infinitesimal mean vector of the matrix B and the vector γ by means of maximum likelihood, in the former diffusion case, the parameters leads to the log normal process with they conditional density. Let $X_{t_1}=x_1,...,x_{t_n}=x_n$ be a discrete sampling of the process in times $t_1,...,t_n$.
with

$$X_{t_\alpha}=(X_{t_\alpha,1},....,X_{t_\alpha,k})'$$

and,

$$x_\alpha=(x_{\alpha 1},...,x_{\alpha k})'$$

Under the condition $P[X_{t_1}=x_1]=1$, the multidimensional likelihood function for a sample of size n will be,

$$L(x_1,...,x_n;\beta,B,\gamma)=(2\pi)^{-\frac{(n-1)}{2}k}|B|^{-\frac{(n-1)}{2}}\prod_{\alpha=2}^{n}\left(\prod_{i=1}^{k}(x_{\alpha i}-\gamma_i)^{-1}\right)$$

$$\times(t_\alpha-t_{\alpha-1})^{-\frac{k}{2}}\exp\left(-\frac{1}{2}(\upsilon_\alpha-\beta u_\alpha)'B^{-1}(\upsilon_\alpha-\beta u_\alpha)\right).$$

Where

$$u_\alpha=(t_\alpha-t_{\alpha-1})^{-\frac{1}{2}}\left(t_\alpha-t_{\alpha-1},\int_{t_{\alpha-1}}^{t_\alpha}F_1(\tau)d\tau,...,\int_{t_{\alpha-1}}^{t_\alpha}F_q(\tau)d\tau\right)'$$

And

$$\upsilon_\alpha=(t_\alpha-t_{\alpha-1})^{-\frac{1}{2}}(ln(x_\alpha-\gamma)-ln(x_{\alpha-1}-\gamma))$$

©2011-2014. All rights reserved.

A. Remark

As γ leads to $x^{(1)} = (x_1^{(1)}, ..., x_k^{(1)})'$, Where ;

$$x_i^{(1)} = inf_{1\leq\alpha\leq n}(x_{\alpha i})$$

for $i = 1, ..., k$, the likelihood function approaches to infinity.
As an alternative to global maximum-likelihood estimators (M.L.E), we might equate partial derivatives of the log-likelihood function to zero, as was done by Cohen[5], Cohen and Witten[6], Harter and Moore[13], to obtain local minimum likelihood estimators(L.M.L.E). As we substitute x_α by υ_α, we obtain:

$$L(\upsilon_2, \ldots, \upsilon_n; \beta, B, \gamma) = (2\pi)^{-\frac{(n-1)}{2}k}|B|^{-\frac{(n-1)}{2}} \prod_{\alpha=2}^{n}$$

$$\exp\left(-\frac{1}{2}(\upsilon_\alpha - \beta u_\alpha)' B^{-1}(\upsilon_\alpha - \beta u_\alpha)\right)$$

The log-likelihood function is

$$[L(\upsilon_2, ..., \upsilon_n; \beta, B, \gamma)] = -\frac{(n-1)}{2}k\ln(2\pi) - \frac{(n-1)}{2}\ln(|B|)$$

$$-\frac{1}{2}tr\left\{\sum_{\alpha=2}^{n}(\upsilon_\alpha - \beta u_\alpha)' B^{-1}(\upsilon_\alpha - \beta u_\alpha)\right\}$$

Using Well-Known results of matrix differential calculations, see the Magnus and Neudecker[17], Rogers [20] and Nel[19]), the differential of the log-likelihood function is:

$$d\ln[L(\upsilon_2, ..., \upsilon_n; \beta, B, \gamma)] = -\frac{(n-1)}{2}tr[B^{-1}(dB)]$$

$$-\frac{(n-1)}{2}tr\sum_{\alpha=2}^{n}\left[-(\upsilon_\alpha - \beta u_\alpha)' B^{-1}(dB)B^{-1}(\upsilon_\alpha - \beta u_\alpha) + \right.$$

$$-2(\upsilon_\alpha - \beta u_\alpha)' B^{-1}(d\beta)u_\alpha - 2(\upsilon_\alpha - \beta u_\alpha)' B^{-1}W_\alpha\alpha(d\gamma)$$

Where

$$W_\alpha = -\frac{d\upsilon_\alpha}{d\gamma} = (t_\alpha - t_{\alpha-1})^{-\frac{1}{2}}$$

$$\begin{pmatrix} \frac{1}{x_{\alpha 1}-\gamma_1} - \frac{1}{x_{\alpha-1}-\gamma_1} & 0 & \cdots & & 0 \\ 0 & & \ddots & \ddots & 0 \\ 0 & & \ddots & \ddots & 0 \\ 0 & & \cdots & 0 & \frac{1}{x_{\alpha k}-\gamma_k} - \frac{1}{x_{\alpha-k}-\gamma_k} \end{pmatrix}$$

$$d\ln[L(\upsilon_2, ..., \upsilon_n; \beta, B, \gamma)] = \frac{1}{2}tr\left\{\left[\sum_{\alpha=2}^{n} B^{-1}(\upsilon_\alpha - \beta u_\alpha)'\right.\right.$$

$$(\upsilon_\alpha - \beta u_\alpha) - I_k \times B^{-1}(dB) + 2\sum_{\alpha=2}^{n}(d\beta)u_\alpha(\upsilon_\alpha - \beta u_\alpha)' B^{-1} +$$

$$\sum_{\alpha=2}^{n}(\upsilon_\alpha - \beta u_\alpha)' B^{-1}W_\alpha(d\gamma)$$

$$d\ln[L(\upsilon_2, ..., \upsilon_n; \beta, B, \gamma)] = \frac{1}{2}Vec\left[B^{-1}\sum_{\alpha=2}^{n}((\upsilon_\alpha - \beta u_\alpha)'\right.$$

$$(\upsilon_\alpha - \beta u_\alpha)B^{-1} - I_k)' dVec(B) + Vec\left[B^{-1}\sum_{\alpha=2}^{n}(\upsilon_\alpha - \beta u_\alpha)u'_\alpha\right]'$$

$$dVec(\beta)\left[\sum_{\alpha=2}^{n}W_\alpha B^{-1}(\upsilon_\alpha - \beta u_\alpha)\right]d\gamma$$

On differentiating the log-likelihood function and equating to zero, we obtain the L.M.L estimating equations:

$$\hat{B}^{-1}\sum_{\alpha=2}^{n}(\hat{\upsilon}_\alpha - \hat{\beta}u_\alpha)u'_\alpha = 0$$

$$\hat{B}^{-1}\sum_{\alpha=2}^{n}\left\{[(\hat{\upsilon}_\alpha - \hat{\beta}u_\alpha)(\hat{\upsilon}_\alpha - \hat{\beta}u_\alpha)'\hat{B}^{-1}] - I_k\right\} = 0$$

$$\sum_{\alpha=2}^{n}\hat{W}_\alpha\hat{B}^{-1}(\hat{\upsilon}_\alpha - \hat{\beta}u_\alpha) = 0$$

$$\hat{\beta} = \sum_{\alpha=2}^{n}\hat{\upsilon}_\alpha u'_\alpha\left[\sum_{\alpha=2}^{n}u_\alpha u'_\alpha\right]^{-1}$$

$$\hat{B} = \frac{1}{n-1}\sum_{\alpha=2}^{n}(\hat{\upsilon}_\alpha - \hat{\beta}u_\alpha)(\hat{\upsilon}_\alpha - \hat{\beta}u_\alpha)'$$

$$\sum_{\alpha=2}^{n}\hat{W}_\alpha\hat{B}^{-1}(\hat{\upsilon}_\alpha - \hat{\beta}u_\alpha) = 0$$

Where

$$u_\alpha = (t_\alpha - t_{\alpha-1})^{-\frac{1}{2}}\left(t_\alpha - t_{\alpha-1}, \int_{t_{\alpha-1}}^{t_\alpha}F_1(\tau)d\tau, \ldots, \int_{t_{\alpha-1}}^{t_\alpha}F_q(\tau)d\tau\right)'$$

$$\hat{\upsilon}_\alpha = (t_\alpha - t_{\alpha-1})^{-\frac{1}{2}}(ln(x_\alpha - \hat{\gamma}) - (ln(x_{\alpha-1} - \hat{\gamma}))$$

$$\hat{W}_\alpha = (t_\alpha - t_{\alpha-1})^{-\frac{1}{2}}$$

$$\begin{pmatrix} \frac{1}{x_{\alpha 1}-\gamma_1} - \frac{1}{x_{\alpha-1}-\gamma_1} & 0 & \cdots & & 0 \\ 0 & & \ddots & \ddots & 0 \\ 0 & & \ddots & \ddots & 0 \\ 0 & & \cdots & 0 & \frac{1}{x_{\alpha k}-\gamma_k} - \frac{1}{x_{\alpha-k}-\gamma_k} \end{pmatrix}$$

And $\hat{\gamma}_i < x_i^{(1)} = inf_{1\leq\alpha}(x_{\alpha i})$ for $i = 1, \ldots, k$
Equation (1) may be solved iteratively for $\hat{\gamma}$, so we will consider the following system equations

$$\hat{\beta}(\gamma) = \sum_{\alpha=2}^{n}\upsilon_\alpha u'_\alpha\left[\sum_{\alpha=2}^{n}\upsilon_\alpha u'_\alpha\right]^{-1}$$

$$\hat{\beta}(\gamma) = \frac{1}{n-1}\sum_{\alpha=2}^{n}(\upsilon_\alpha - \hat{\beta}(\gamma)u_\alpha)(\upsilon_\alpha - \hat{\beta}(\gamma)u_\alpha)'$$

$$\lambda(\hat{\beta}(\gamma), \hat{\beta}(\gamma), \gamma) = \sum_{\alpha=2}^{n}W_\alpha\hat{B}^{-1}(\gamma)(\upsilon_\alpha - \hat{\beta}(\gamma)u_\alpha)$$

With $\upsilon_\alpha = (t_\alpha - t_{\alpha-1})^{-\frac{1}{2}}(ln(x_\alpha - \gamma) - (ln(x_{\alpha-1} - \gamma))$.

Finally, standard iterative with fixed derivatives,Newton-Raphson, and Fisher Scoring methods, are satisfactory for solving the equations system.

©2011-2014. All rights reserved.

REFERENCES

[1] A. Arbai, "Procesos estoclognormales triparamtricos", Tesis Doctoral Univ Granada, 1994

[2] I.V. Bassawa and R.Prakasa, " Statistical inference for Stochastic Process," Academic Press, 1980.

[3] B.M Brown and J.I.Hewitt, "Asymptotic likelihood theory for diffusion process," Theory Probab Its Appl, (1986), 2 373-377.

[4] F.Calitz, "Maximum likelihood estimation of three-parameter log normal distribution a reconsideration," Austral, J.Statist, (1973), 15 185 190

[5] A.C.Cohen, "Estimating parameters of logarithmic-normal distributions by maximum likelihood," J.Amer.Statist.Assoc, (1951), 46 206 212.

[6] A.C.Cohen and B.J.Witten, "Estimation in the three-parameter log normal distribution," J.Amer.Stat.Assoc, (1980), 75 399 404.

[7] A.C.Cohen and B.J Witten and Y.Ding, " Modified moment estimation for the three-parameter log normal distribution," J.Qual.Tech, (1985), 17 92 99

[8] S.L.Crawford, "An application of the Laplace method to finite mixture distributions, American Statistical Association," (1994)

[9] F.Giesbretcht and O.Kempthorne, "Maximum likelihood estimation in the three parameter log normal distribution," J.Roy.Statis, (1976), 29 58 68.

[10] R.Gutiérrez and A.Hermoso and M. Molina, "Test de rezón de verosimilitudes para contrastes de hipó tesis sobre el coeficiente tendencia en procesos logaritmico-normales multidimensionales con factores exógenos," Cuadernos de Estadistica, (1984), 7 8 129 136.

[11] R.Gutiérrez and J.M.Angulo and A.González and R.Pérez ,"Inference in log normal multidimensional diffusion process with exogenous factors: application to modeling in economics. Applied Stochastic Models and Data Analysis," (1991), 7(4) 293-316

[12] D.A Griffiths , "Interval estimation for the three-parameter log-normal distribution via the likelihood function," Applied Statics, (1980), 29 85-68.

[13] H.L.Harter and A.H.Moore, " Local-maximum likelihood estimation of the parameters of three-parameter log normal populations from complete and censored samples," J.Amer.Statics.Assoc, (1966), 61 842-851

[14] A.Hermoso, "Test de hipótesis sobre el coeffciente tendencia de un process de diffusión multidimensional. Applicatión al proceso logaritmico normal con factores exogenós," Tesis Doctoral Univ. Granada, (1984)

[15] B.M.Hill, "The three parameter log normal distribution and bayesian analysis of a point-source epidemic," J.Amer.Statist.Assoc, (1963), 58 72-84.

[16] J.A.Lambert"Estimation for the three-parameter log normal distribution ," Australian journal of statistics,(1964), 6,29-32.

[17] J.R.Magnus and H.Neudecker, "Matrix differential calculus with applications in statistics and econometrics," John Whiley Sons, (1988)

[18] M.Molina, "Estimación del coeficiente tendencia de un proceso de defusión multidimentional: applicatión al proceso logaritmico normal con factores exógenos," Tesis Doctoral, Univ.Granada, (1984)

[19] D.G.Nel, "On matrix differentiation in statistics," South African Statistics Journal,(1980), 14 137 193.

[20] S.Rogers, "Matrix Derivatives Lecture notes in statistics," Vol2, (1980

[21] Royston, P "Estimation reference ranges and goodness of fit for the three-parameter log normal distribution," Statistics in medicine, (1992), 11, 897-912.

[22] F.Torres, "Aportationes al estudio de difusiones estocasticas no homogeneas," Tesis Doctoral Univ.Granada, (1993)

[23] G.Tintner and R.Narayanan, "A multidimensional stochastic process for the explonation of economic development," Metrika, (1966), 11 85-90.

[24] G.Tintner and J.K. Sengupta, "Stochastic Economic," Academic Press, (1972)

[25] G.Tintner and G.L.Gomez, "Applicatión de los procesos de difusión a la planificación del desarrallo económico," Un ejemplo ilustrativo(Colombia).Trabajos de Estadistica, (1979), 30(2)-33-55

[26] S.D.Wicksell, "On logarithmic correlation with an application to the distribution of ages first marriage," Meddelande Fran lunds tronomiska Observatorium,84, (1917)