Lognormal diffusion process with polynomial exogenous factors

Ramón Gutiérrez Jáimez, Patricia Román Román, Nuria Rico Castro, Desirée Romero Molina, Francisco Torres Ruiz
University of Granada, Department of Statistics and Operation Research
Avda. Fuentenueva s/n
Granada (18071),Spain
rgjaimez@ugr.es, nrico@ugr.es, proman@ugr.es, deromero@ugr.es, fdeasis@ugr.es

In this paper we propose the use of polynomial exogenous factors in the lognormal diffusion process as an approximation in the case that no external information about the process is available or the functional form of the exogenous factors is unknown.

Introduction

The use of diffusion processes that include exogenous factors in their trend is common in many application fields. The reason of its application is the usual presence of deviations of the observed data with respect to the trend of some known homogenous diffusion process, in some time intervals. These factors are time dependent functions that allow, on one hand, a best fit to the data in the situations mentioned above and, on the other hand, an external control to the process behaviour.

One of the mentioned kind of processes is the lognormal diffusion process with exogenous factors. This process is defined as a diffusion process with infinitesimal moments $A_1(x,t) = xh(t)$ and $A_2(x,t) = \sigma^2 x^2$, where $\sigma^2 > 0$ and h is a continuous function in $[t_0, T]$ containing the external information to the process. Frequently, there are more of one external information sources, so it is usual to take h as a linear combination of continuous functions F_i, $i = 1, \ldots, q$ (from now, factors).

Some topics, as the inference and first-passage-time through varying boundaries, are already studied. It is important to note that for the inferential process is not necessary the functional form of the exogenous factors but the value of their integrals between two time values included in the interval $[t_0, T]$ (analogous to some interpolation problems) whereas the probability density function of the first-passage-time of the process through a boundary includes the factors explicitly (Gutiérrez et al. (1999)).

However, it is usual in practice that the functional form of the exogenous factors is not available or, indeed, the external influences are unknown. This paper shows a possible approximation in such situations by using polynomial exogenous factors, that is, in this case $h(t) = \sum_{j=0}^{k} \beta_j^{k)} P_j(t)$, where P_j is a j-degree polynomial, $j = 1, \ldots, k$ and $P_0(t) = 1$.

Some aspects about the inference on the model

Let $\{X(t) : t_0 \leq t \leq T\}$ be the lognormal diffusion process with polynomial exogenous factors. Some of the main characteristics of the process can be expressed in terms of a kind of parametric functions, concretely $\theta(C, \mathbf{A}(t,s), B(t,s), l) = C \exp\left(\mathbf{A}'(t,s)\mathbf{a}_k + B(t,s)\sigma_k^l\right)$, with $\mathbf{a}_k = (\beta_0^{k)} - \frac{\sigma_k^2}{2}, \beta_1^{k)}, \ldots, \beta_k^{k)})'$, $C > 0$, $l \in N$, $\mathbf{A}(t,s) \in R^{k+1}$ and $B(t,s) \in R$. Two of these functions are the trend and conditioned trend that can be written, respectively, as $m(t) = \mathrm{E}[X(t)] = \theta\left(\mathrm{E}[X(t_0)], \bar{\mathbf{u}}^{k)}(t), \frac{t-t_0}{2}, 2\right)$ and $m(t|s) = \mathrm{E}[X(t)|X(s) = x_s] = \theta\left(x_s, \bar{\mathbf{u}}^{k)}(t,s), \frac{t-s}{2}, 2\right)$, where $\bar{\mathbf{u}}^{k)}(t,s) = \left(t-s, \int_s^t P_1(\tau)d\tau, \ldots, \int_s^t P_k(\tau)d\tau\right)'$ and $\bar{\mathbf{u}}^{k)}(t) = \bar{\mathbf{u}}^{k)}(t, t_0)$.

The ML estimators of the parameters $\mathbf{a}_k$ and σ_k^2 are $\widehat{\mathbf{a}}_k = \mathbf{V}_k\mathbf{v}$ and $\widehat{\sigma}_k^2 = \frac{1}{n-1}\mathbf{v}'\mathbf{H}_k\mathbf{v}$, where $\mathbf{v} = (v_2, \ldots, v_n)'$, $\mathbf{V}_k = (\mathbf{U}_k\mathbf{U}_k')^{-1}\mathbf{U}_k$ and $\mathbf{H}_k = \mathbf{I}_{n-1} - \mathbf{U}_k'\mathbf{V}_k$, being $\mathbf{U}_k$ the matrix, whose rank is assumed to be $k+1$, given by $\mathbf{U}_k = (\mathbf{u}_2^{k)}, \ldots, \mathbf{u}_n^{k)})$ with $\mathbf{u}_i^{k)} = (t_i - t_{i-1})^{-1/2}\bar{\mathbf{u}}^{k)}(t_i, t_{i-1})$. From these estimators one can obtain the corresponding to the θ functions. The ML estimator is immediate whereas the UMVU estimator, for $l = 2$, is $Ce^{\mathbf{A}'(t,s)\widehat{\mathbf{a}}_k}\,{}_0F_1\left(\frac{n-k-2}{2}; \frac{(n-1)\left[B(t,s)-\frac{1}{2}A_{t,s}^{U_k}\right]}{2}\widehat{\sigma}_k^2\right)$ (Gutiérrez et al. (2001)), with $A_{t,s}^{U_k} = \mathbf{A}'(t,s)(\mathbf{U}_k\mathbf{U}_k')^{-1}\mathbf{A}(t,s)$.

One question that can be formulated in this model is what is the optimal degree of the polynomial, that is, how many exogenous factors must be taken. The answer is not immediate because it will depend on the data. But one can take a similar strategy to that used in polynomial regression. For example, a forward procedure can be followed by introducing successively polynomials in the model. In such situation, and after introducing P_{k+1}, the information about the exogenous factors is in the matrix $\mathbf{U}_{k+1} = (\mathbf{U}_k'|\mathbf{d}_{k+1})'$, where the $(i-1)$-th component of the d_{k+1} vector is given by $(t_i - t_{i-1})^{-\frac{1}{2}}\int_{t_{i-1}}^{t_i} P_{k+1}(\tau)d\tau$, $i = 2, \ldots, n$. Taking into account the expression of the inverse of a partitioned matrix, and denoting $e_k = \mathbf{d}_{k+1}'\mathbf{H}_k\mathbf{d}_{k+1}$, one obtain

$$\widehat{\mathbf{a}}_{k+1} = \frac{1}{e_k}\left[\begin{pmatrix} e_k\mathbf{I}_{k+1} + \mathbf{V}_k\mathbf{d}_{k+1}\mathbf{d}_{k+1}'\mathbf{U}_k' \\ -\mathbf{d}_{k+1}'\mathbf{U}_k' \end{pmatrix}\widehat{\mathbf{a}}_k + \begin{pmatrix} -\mathbf{V}_k\mathbf{d}_{k+1} \\ 1 \end{pmatrix}\mathbf{d}_{k+1}'\mathbf{v}\right],$$

$\widehat{\sigma}_{k+1}^2 = \widehat{\sigma}_k^2 - \dfrac{\mathbf{v}'\mathbf{H}_k\mathbf{d}_{k+1}\mathbf{d}_{k+1}'\mathbf{H}_k\mathbf{v}}{e_k}$ and $\bar{\mathbf{u}}^{k+1)'}(t,s)(\mathbf{U}_{k+1}\mathbf{U}_{k+1}')^{-1}\bar{\mathbf{u}}^{k+1)}(t,s) =$
$\bar{\mathbf{u}}^{k)'}(t,s)(\mathbf{U}_k\mathbf{U}_k')^{-1}\bar{\mathbf{u}}^{k)}(t,s) + \left[\bar{\mathbf{u}}^{k)'}(t,s)\mathbf{V}_k\mathbf{d}_{k+1} - \int_s^t P_{k+1}(\tau)d\tau\right]^2 / e_k$.

In this development it is necessary to note that the $\mathbf{U}_k$ matrix contains powers of the t_i, so the explained procedure is more reliable for low degrees because of the conditioning of the matrix $\mathbf{U}_k\mathbf{U}_k'$. In this sense, the above recursive expressions can ameliorate the accuracy of the calculations. Nevertheless, for some instances, orthogonal polynomials or more smoothed functions (like splines) will be preferred.

Examples

We illustrate this situation with two examples, the first based on a simulated process and the second on real data of the Gross National Product of Spain, being the choice of the exogenous polynomial factors and the fit of the theoretical model to the data the main interest of these examples.

REFERENCES

Gutiérrez, R., Román, P. and Torres, F. (1999). Inference and first-passage time for the lognormal diffusion process with exogenous factors: application to modelling in economics. Applied Stochastic Models in Business and Industry, 15, 325-332.

Gutiérrez, R., Román, P. and Torres, F. (2001). Inference on some parametric functions in the univariate lognormal diffusion process with exogenous factors. Test, 10(2), 357-373.

RÉSUMÉ

Dans ce papier nous proposons l'utilisation des facteurs exogènes polynomiales dans le processus de diffusion lognormale comme une approximation soit dans le cas où on n'a pas d'information externe concernant le processus, soit quand la forme fonctionnelle des facteurs exogènes est inconnue. Cette situation est illustrée avec deux exemples, le premier basé sur un processus simulé et le deuxième sur des datas réels obtenus à partir du Produit Intérieure Brut espagnol. L'intérêt majeur de ces exemples est le choix des facteurs exogènes polynomiales et comment le modèle théorique s'ajuste aux datas.