The Korean Journal of Applied Statistics (2018) DOI: http://dx.doi.org/10.5351/KJAS.2018.31.6.801
31(6), 801–810

Random effect models for simple diffusions

Eun-Kyung Lee[a] · In Suk Lee[b] · Yoon Dong Lee[b,1]

[a]Department of Statistics, Ewha Womans University; [b]Business School, Sogang University

(Received October 29, 2018; Revised November 7, 2018; Accepted November 7, 2018)

Abstract

Diffusion is a random process used to model financial and physical phenomena. When we construct statistical models for repeatedly observed diffusion processes, the idea of random effects needs to be considered. In this research, we introduce random parameters for an Ornstein-Uhlenbeck diffusion model and geometric Brownian motion diffusion model. In order to apply the maximum likelihood estimation method, we tried to build likelihoods in closed-forms, by assuming appropriate distributions for random effects. We applied the random effect models to data consisting of Dow Jones Industrial Average indices recorded daily over 27 years from 1991 to 2017.

Keywords: diffusion, random effects, OU model, GBM model

1. 서론

입자의 운동모형에서 입자의 위치를 확률적인 모형으로 설명하는 경우, 불확실성(randomness)과 연속성(continuity)을 만족하여야 하고, 이런 경우 입자의 위치는 시각 $t\,(t \geq 0)$에 대하여 연속적인 경로를 갖는 브라운운동 w_t를 이용한 확률편미분방정식을 통하여 표현된다.

$$dy_t = \mu_t dt + \sigma_t dw_t. \tag{1.1}$$

확산 y_t는 추세계수 μ_t와 확산계수 σ_t를 갖는 확률편미분방정식 식 (1.1)의 해로 정의되는 확률과정이다. 입자의 운동, 금융자산의 가격변동, 투여된 약물 농도의 변화 등이 모두 확산(diffusion)의 예이다.

다양한 변동특성을 갖는 금융자산의 가격변동을 확산의 개념을 이용하여 설명하기 위하여 다양한 형태의 모형이 도입되었다. 금융위기 상황에서 자산 가격의 급격한 변동을 설명하기 위한 방법으로, 경로가 불연속적인 특성을 가질 수 있도록 점프(jump) 확률과정을 포함하는 일반화된 모형이 고려되기도 하고 (Applebaum, 2004), 기본적인 확산모형이 보다 높은 변동성을 포용할 수 있도록 여러 확률과정을 결합하여 사용하는 모형들이 개발되었다 (Heston, 1993).

짧은 시간 동안의 입자의 운동이나 자산의 가격변동을 설명하는 데 있어서 경로의 연속성은 마땅히 만족되어야 할 중요한 조건이다. 그러나 장기간 동안 관찰된 입자의 운동이나, 금융자산의 가격변동을 동

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2015R1D1A1A 01056790).
[1]Corresponding author: Business School, Sogang University, 35 Baek Bum Ro, Seoul 04107, Korea.
E-mail: widylee@sogang.ac.kr

일한 하나의 확산모형으로 설명해야 한다는 것은 지나친 제약이다. 공기 중 입자의 움직임에서 순간적으로 바람의 방향이 달라지는 경우, 혹은 예상치 못한 금융정책의 발표로 금융자산의 가격이 순간적으로 변동하는 경우와 같이, 장시간 동안의 입자의 움직임이나 금융자산의 변화에서 변동 양태의 연속성을 가정하기 어려운 사건들이 발생할 수 있다고 보아야 한다.

입자 운동에서의 물리적 환경이나 주가의 변동에서 금융 정책의 변화와 같이 모형의 변동요인이 매우 드물게 혹은 가끔씩 일어난다고 가정하는 경우에는 시간에 따라 확산모형의 모수가 변화하는 형태로 모형을 일반화 하여 설정하는 방법도 생각할 수 있으나, 확산현상 중에는 관측되는 확률과정의 모수의 변동요인이 상시적으로 발생한다고 보아야 하는 경우들이 더욱 일반적이다. 공기 중 입자의 운동을 설명하는 확산모형에서 바람의 변화는 매우 중요한 변동 요인이지만 그 변화는 수시로 발생할 수 있고, 바람의 변화는 통제되거나 예측이 가능하지 않은 경우가 더 일반적이다. 또 주식가격의 변동은, 매일 매일 형성되는 시장의 투자 분위기에 따라 크게 영향을 받고, 시장의 분위기는 예측 가능하지 않은 경우들이 많다.

이와 같이 예측 불가능한 다양한 요인에 의하여 모형의 특성이 수시로 변동되는 확산모형을 고려하기 위해서는, 추세계수 μ_t와 확산계수 σ_t에 확률적 특성을 갖는 값을 도입하는 확률효과모형(random effect model)을 고려하여야 한다. 본 논문에서는 매우 단순한 확산과정들을 가정하고 그 계수들이 불확실성을 갖는 것으로 가정하고, 즉 모형모수가 확률효과인 모형을 가정하고, 이에 대한 최대우도추정법을 적용하기 위하여 우도함수를 구성하는 방법을 살펴본다. 선형 혹은 비선형 형태의 회귀모형에서 확률효과를 도입하는 경우에서 최대우도추정법에 의한 모수 추정에 대한 연구로는 Laird와 Ware (1982)와 Pinheiro와 Bates (2000)가 있다.

확산모형에 확률효과를 도입하는 연구는 선행연구가 많지 않고, 최근 일부 시도가 이루어지고 있으나, 제한적인 조건 하에서 근사적 방법 등을 이용하고 있다 (Picchini 등, 2010; Delattre, 2013). 확산모형의 확률효과 모형에 대한 연구가 활발하게 진행되지 못하는 기본적인 이유는, 일부 단순한 모형을 제외한 대부분의 확산모형에서 전이확률밀도함수가 닫힌 형태(closed-form)로 구할 수 있는 경우가 없기 때문이다.

확산모형에 대한 전이확률밀도함수를 구하는 과정은 Ait-Sahalia (2002)에 의하여 정리된 바 있고, Lee 등 (2014)는 이를 개선하여 매우 정확한 정도로 전이확률밀도함수를 구하는 방법을 제시하였다. 본 논문에서는 일반적인 확산모형에 대하여 Lee 등 (2014)에서 제시한 방법을 도입하기에 앞서, 전이확률밀도함수의 수식이 알려져 있는 단순한 확산모형들에, 적절한 형태의 분포를 갖는 확률효과를 도입하여, 수치적 방법 등에 의존하지 않고, 수식에 의한 적분이 가능하도록 하여, 주변분포를 구해보고, 그 형태가 어떻게 구해지는 지를 살펴보게 된다.

본 논문에서는, 금융현상의 설명에 필수적인 확산모형인 Geometric Brownian Motion (GBM) 모형을 중심으로, 확률효과모형을 설정하고, 실제 자료 분석에 이용한다. 또 기본적인 확산모형인 Ornstein-Uhlenbeck (OU) 모형에 확률효과를 도입하는 경우에 대하여도, 함께 살펴보게 된다.

본 논문의 제2절에서는 기본적인 확산모형인 OU 모형과 GBM 모형에 대하여 살펴보고, 제3절에서는 이를 확률효과모형으로 확장하여, 그 우도함수를 구하는 방법을 설명한다. 제4절에서는 실제 자료에 확률효과가 있는 GBM 모형을 적용하여 모수 추정 과정을 살펴보게 된다.

2. 확산모형

2.1. 확산모형의 설정

본 연구에서는 입자의 운동을 설명하는 경우, 금융자산의 변동을 설명하는 경우, 각 경우에 적용되는 확

산모형 중 가장 단순하고 기초적인 형태의 확산모형을 가정하고, 이후 각 모형의 모수에 대하여 확률효과를 도입하는 과정을 살펴본다.

$$\begin{aligned} \text{OU 모형} : dy_t &= \alpha(\beta - y_t)dt + \sqrt{2\alpha\eta}\, dw_t, \quad \alpha, \eta > 0, \\ \text{GBM 모형} : ds_t &= \beta s_t dt + \sqrt{2\eta}\, s_t dw_t, \quad \eta > 0. \end{aligned} \tag{2.1}$$

OU 모형은, 입자의 움직임을 설명하는 기본적인 확산모형이고, GBM 모형은 금융자산의 가격변동을 설명하는 기본적인 모형이다. GBM 모형에서 $y_t = \log s_t$와 같이 변환하면 다음과 같다.

$$dy_t = (\beta - \eta)dt + \sqrt{2\eta}\, dw_t. \tag{2.2}$$

본 논문에서는 확산에 대한 확률효과모형에서의 모수 추정의 문제를 살펴보게 되는데, GBM 확률과정 s_t에 대한 관측값이 주어지는 것은 $y_t = \log s_t$에 대한 관측값이 주어지는 것과 동일하고, 확률과정에 대한 관측값이 주어지는 경우, 모수 추정의 문제에서는 모수들 사이의 관계를 어떻게 설정할 것인가가 모형설정의 핵심적 사항이므로, 본 논문에서는 s_t 대신 $y_t = \log s_t$에 대한 식 (2.2)를 이용하여 GBM 모형에 대한 추정 문제를 다룬다.

OU 모형에서 모수 α는 시간 t와의 곱으로 개입되는 모수로, 시간의 단위의 역수 혹은 확률과정의 시간에 대한 반응속도를 결정하는 모수다. 위 식 (2.1)과 (2.2)에서 모수 α, β, η를 관례에 따라 각각 속도모수, 평균모수, 확산모수라고 부르기로 한다.

2.2. 관측된 확산모형과 전이밀도함수

확산 확률과정 y_t에서 시간 t는 연속적인 값이다. 즉, 확산은 모형적 측면에서는 시간에 대하여 연속적인 특성을 갖는다. 그러나 이를 관측하는 경우, 매우 조밀한 시간 간격으로 관측이 이루어진다고 가정하더라도, 관측이란 결국은 이산적인 시간간격으로 이루어질 수밖에 없다.

확산 y_t가 시간간격 δ로 관측이 이루어진 경우, 즉, $t = 0$인 시점에서 y_0가 관측되고, $t = \delta$인 시점에서 y_δ를 관측하는 경우, 관측 값 y_δ의 확률적 특성은 y_0가 주어진 경우에 대한 조건부분포로 주어지고, 확률밀도함수 $f(y_\delta|y_0)$로 표현된다. 이때의 확률밀도함수를 전이확률밀도함수라 한다.

다양한 확산모형 중 OU 모형, GBM 모형, CIR 모형의 경우에는 전이확률밀도함수를 닫힌 형태로 표현이 가능하다. 일반적인 확산모형에 대한 전이확률밀도함수는 포커플랑크방정식의 해로 표현된다 (Hurn 등, 2007). Lee 등 (2014)는 전이확률 밀도함수를 허밋다항식을 기저로 하여 전개하는 델타확장법을 완성하였다.

만약 규칙적인 시간간격 δ마다 확산에 대한 관측이 이루어지고, 그 관측값들로 이루어진 관측벡터 $\boldsymbol{y} = (y_0, y_\delta, y_{2\delta}, \ldots, y_{n\delta})$을 얻은 경우, 관측벡터 분포를 표현하는 결합분포의 확률밀도함수는 다음과 같다.

$$f(\boldsymbol{y}) = \prod_{i=1}^{n} f\left(y_{i\delta}|y_{(i-1)\delta}\right).$$

• OU 모형의 결합분포함수:

y_0가 주어진 경우 y_δ의 조건부분포는, $u = (1 - e^{-\alpha\delta})$이고 $v = (1 - e^{-2\alpha\delta})$라고 할 때, 평균이 $(1 - u)y_0 + u\beta$이고, 분산이 ηv인 정규분포의 형태로 주어지므로, 즉,

$$y_\delta \sim \text{Normal}\left((1-u)y_0 + u\beta,\ \eta v\right)$$

이므로, 결합분포의 밀도함수는

$$f(\boldsymbol{y}) = (2\pi\eta v)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}\frac{1}{\eta v}\sum_{i=1}^{n}(c_i - u\beta)^2\right\}$$

이고, 여기서 $c_i = y_{i\delta} - (1-u)y_{(i-1)\delta}$, $i = 1, 2, \ldots, n$이다.

• GBM 모형의 결합분포함수:

y_0가 주어진 경우 y_δ의 조건부분포는, 평균이 $y_0 + \delta(\beta - \eta)$이고, 분산이 $2\eta\delta$인 정규분포의 형태로 주어지므로, 즉,

$$y_\delta \sim \text{Normal}\left(y_0 + \delta(\beta - \eta),\, 2\eta\delta\right)$$

이므로, 결합분포는 $d_i = y_{i\delta} - y_{(i-1)\delta}$, $i = 1, 2, \ldots, n$이고, $\tilde{\beta} = \beta - \eta$라고 할 때,

$$f(\boldsymbol{y}) = (2\pi 2\eta\delta)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}\frac{1}{2\eta\delta}\sum_{i=1}^{n}\left(d_i - \delta\tilde{\beta}\right)^2\right\}$$

라고 표현되거나, $b_i = y_{i\delta} - y_{(i-1)\delta} - \delta\beta$, $i = 1, 2, \ldots, n$라고 할 때,

$$f(\boldsymbol{y}) = (2\pi 2\eta\delta)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}\frac{1}{2\eta\delta}\sum_{i=1}^{n}(b_i + \eta\delta)^2\right\}$$

라고 표현된다.

3. 확률효과모형

확산 확률과정 에 대한 관측이 여러 차례 반복되어 여러 개의 관측벡터 $\boldsymbol{y}^{(k)} = (y_0^{(k)}, y_\delta^{(k)}, \ldots, y_{n\delta}^{(k)})$, $k = 1, 2, \ldots, m$가 얻어지고, 각 경우에서 확산모형의 모수가 확률적으로 변화되는 확률효과인 경우를 가정하자. 일반적인 접근법으로는 각 모형에서 사용되는 모수 α, β, η를 동시에 확률효과로 확장하는 경우를 고려할 수도 있겠으나, 복수의 확률효과를 고려하는 경우, 우도함수를 구하기 위하여 다중적분을 계산하여야 하는 어려움이 있다. 본 논문에서는 다중적분의 어려움을 피하기 위하여 모수 α, β, η를 각각 별개의 확률효과로 가정하는 경우에 대하여, 최대우도법으로 모형모수를 추정하기 위하여 우도를 구성하는 방법을 살펴본다.

확률효과모형에서, 확률효과를 $\xi^{(k)}$라고 하고 모형 모수를 θ라 하면, 모형모수에 대한 우도함수 $l(\theta)$는, 관측벡터에 대한 조건부 결합분포 $f(\boldsymbol{y}^{(k)}|\xi^{(k)})$를 확률효과의 분포로 적분하여, 주변분포의 밀도함수 $h(\boldsymbol{y}^{(k)}|\theta)$를 구하고 이로부터 다음과 같이 구해진다.

$$h\left(\boldsymbol{y}^{(k)}|\theta\right) = \int f\left(\boldsymbol{y}^{(k)}|\xi^{(k)}\right) g\left(\xi^{(k)}|\theta\right) d\xi^{(k)},$$

$$l(\theta) = \sum_{k=1}^{m} \log h\left(\boldsymbol{y}^{(k)}|\theta\right).$$

다음 소절에서는, OU 모형과 GBM 모형 각 경우에 대하여 관측벡터에 대한 주변분포의 밀도함수 $h(\cdot)$를 구하는 과정을 구체적으로 살펴보게 된다. 모형으로 고려하는 확률효과의 분포 $g(\cdot)$는, 확률효과가 만족해야 하는 제약조건을 잘 만족하고(예: 비음 조건), 분포모수를 통하여 분포의 평균과 분산 등 확률효과가 갖는 분포의 특성을 자유롭게 표현할 수 있는 분포 중에서, 적분을 통하여 관측벡터의 주변분포 $h(\cdot)$가 잘 구해질 수 있고, 그 수리적 형태가 다루기 쉬운 형태의 분포가 되도록 하였다.

3.1. OU 모형의 경우

• 평균모수 확률효과모형:

각 관측벡터 $\boldsymbol{y}^{(k)} = (y_0^{(k)}, y_\delta^{(k)}, \ldots, y_{n\delta}^{(k)})$에 대한 조건부결합분포의 밀도함수와 확률효과 $\beta^{(k)}$의 분포는 다음과 같다.

$$f\left(\boldsymbol{y}^{(k)}|\beta^{(k)}\right) = (2\pi\eta v)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}\frac{1}{\eta v}\sum_{i=1}^{n}\left(c_i^{(k)} - u\beta^{(k)}\right)^2\right\}, \tag{3.1}$$

$$\beta^{(k)} \sim \mathrm{Normal}(\beta, \tau), \tag{3.2}$$

여기서 $u = (1 - e^{-\alpha\delta})$, $v = (1 - e^{-2\alpha\delta})$이고, $c_i^{(k)} = y_{i\delta}^{(k)} - (1-u)y_{(i-1)\delta}^{(k)}$, $i = 1, 2, \ldots, n$, $k = 1, 2, \ldots, m$이다. 즉, 확률효과 $\beta^{(k)}$의 확률밀도함수는 다음과 같다.

$$g\left(\beta^{(k)}|\beta, \tau\right) = (2\pi\tau)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\tau}\left(\beta^{(k)} - \beta\right)^2\right\}.$$

정리 3.1 관측벡터의 조건부결합분포가 식 (3.1)과 같이 주어지는 OU 모형에서, 확률효과 $\beta^{(k)}$의 분포가 식 (3.2)와 같이 주어질 때, 관측벡터 $\boldsymbol{y}^{(k)}$에 대한 주변분포의 밀도함수 h는 다음과 같다.

$$\begin{aligned} h\left(\boldsymbol{y}^{(k)}|\alpha, \beta, \eta, \tau\right) &= (2\pi\eta v)^{-\frac{n}{2}}\left[1 + \frac{u^2}{v}\frac{\tau}{\eta/n}\right]^{-\frac{1}{2}} \\ &\quad \cdot \exp\left\{-\frac{1}{2}\frac{1}{v(\eta/n)}\left(\left[\overline{c^2}^{(k)} - (\bar{c}^{(k)})^2\right] + \left(1 + \frac{u^2}{v}\frac{\tau}{\eta/n}\right)^{-1}\left[\bar{c}^{(k)} - u\beta\right]^2\right)\right\} \end{aligned}$$

여기서, $\bar{c}^{(k)} = (1/n)\sum_{i=1}^{n} c_i^{(k)}$이고 $\overline{c^2}^{(k)} = (1/n)\sum_{i=1}^{n}(c_i^{(k)})^2$이다.

증명: 생략 □

• 확산모수 확률효과모형:

각 관측벡터 $\boldsymbol{y}^{(k)}$에 대한 조건부결합분포의 밀도함수는 식 (3.1)과 동일하고, 확률효과 $\eta^{(k)}$의 분포는 다음과 같다.

$$\left(\eta^{(k)}\right)^{-1} \sim \tau\chi^2(\nu). \tag{3.3}$$

즉, $(\eta^{(k)})^{-1}$의 분포는 규모모수가 τ이고, 자유도가 ν인 카이제곱분포 (규모모수가 2τ이고, 모양모수가 $\nu/2$인 감마분포)이다. 확률효과 $\eta^{(k)}$가 갖는 확률밀도함수 g는 다음과 같이 표현된다.

$$g\left(\eta^{(k)}|\nu, \tau\right) = \frac{1}{\Gamma(\nu/2)(2\tau)^{\frac{\nu}{2}}}\left(\eta^{(k)}\right)^{-\frac{\nu}{2}-1} \exp\left\{-\frac{1}{\eta^{(k)}2\tau}\right\}.$$

정리 3.2 관측벡터의 조건부결합분포가 식 (3.1)과 같이 주어지는 OU 모형에서, 확률효과 $\eta^{(k)}$의 분포가 (3.3)과 같이 주어질 때, 관측벡터 $\boldsymbol{y}^{(k)}$에 대한 주변분포의 밀도함수 h는 다음과 같다.

$$h\left(\boldsymbol{y}^{(k)}|\alpha, \beta, \nu, \tau\right) = \frac{\Gamma((n+\nu)/2)}{(\pi)^{\frac{n}{2}}\Gamma(\nu/2)}\left(\frac{v}{\tau}\right)^{-\frac{n}{2}}\left(1 + \frac{q^{(k)}}{(v/\tau)}\right)^{-\frac{1}{2}(n+\nu)},$$

여기서 $q^{(k)} = \sum_{i=1}^{n}(c_i^{(k)} - u\beta)^2$이다.

증명: 생략 □

3.2. GBM 모형의 경우

- 평균모수 확률효과모형:

각 관측벡터 $\boldsymbol{y}^{(k)} = (y_0^{(k)}, y_\delta^{(k)}, \ldots, y_{n\delta}^{(k)})$에 대한 조건부결합분포의 밀도함수와 확률효과 $\beta^{(k)}$의 분포는 다음과 같다.

$$f\left(\boldsymbol{y}^{(k)}|\beta^{(k)}\right) = (2\pi 2\eta\delta)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}\frac{1}{2\eta\delta}\sum_{i=1}^{n}\left(d_i^{(k)} - \delta\tilde{\beta}^{(k)}\right)^2\right\}, \tag{3.4}$$

$$\beta^{(k)} \sim \text{Normal}(\beta, \tau)\text{이고} \quad \tilde{\beta}^{(k)} \sim \text{Normal}(\tilde{\beta}, \tau) \tag{3.5}$$

이다. 여기서 $\tilde{\beta}^{(k)} = \beta^{(k)} - \eta$, $\tilde{\beta} = \beta - \eta$이고, $d_i^{(k)} = y_{i\delta}^{(k)} - y_{(i-1)\delta}^{(k)}$, $i = 1, 2, \ldots, n$, $k = 1, 2, \ldots, m$이다.

정리 3.3 관측벡터의 조건부결합분포가 식 (3.4)와 같이 주어지는 GBM 모형에서, 확률효과 $\beta^{(k)}$의 분포가 (3.5)와 같이 주어질 때, 관측벡터 $\boldsymbol{y}^{(k)}$에 대한 주변분포의 밀도함수 h는 다음과 같다.

$$\begin{aligned} h\left(\boldsymbol{y}^{(k)}|\beta, \eta, \tau\right) &= (2\pi\eta 2\delta)^{-\frac{n}{2}}\left[1 + \frac{\delta}{2}\frac{\tau}{\eta/n}\right]^{-\frac{1}{2}} \\ &\cdot \exp\left\{-\frac{1}{2}\frac{1}{2\delta(\eta/n)}\left(\left[\overline{d^2}^{(k)} - \left(\bar{d}^{(k)}\right)^2\right] + \left(1 + \frac{\delta}{2}\frac{\tau}{\eta/n}\right)^{-1}\left[\bar{d}^{(k)} - \delta\tilde{\beta}\right]^2\right)\right\} \end{aligned}$$

여기서, $\bar{d}^{(k)} = (1/n)\sum_{i=1}^n d_i^{(k)}$이고 $\overline{d^2}^{(k)} = (1/n)\sum_{i=1}^n (d_i^{(k)})^2$이다.

증명: 생략 □

- 확산모수 확률효과모형:

각 관측벡터 $\boldsymbol{y}^{(k)}$에 대한 조건부결합분포의 밀도함수와 확률효과 $\eta^{(k)}$의 분포는 다음과 같다.

$$f\left(\boldsymbol{y}^{(k)}\right) = \left(2\pi 2\eta^{(k)}\delta\right)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}\frac{1}{2\eta^{(k)}\delta}\sum_{i=1}^{n}\left(b_i^{(k)} + \delta\eta^{(k)}\right)^2\right\}, \tag{3.6}$$

$$\left(\eta^{(k)}\right)^{-1} \sim \tau\chi^2(\nu) \tag{3.7}$$

여기서 $b_i^{(k)} = y_{i\delta}^{(k)} - y_{(i-1)\delta}^{(k)} - \delta\beta$, $i = 1, 2, \ldots, n$, $k = 1, 2, \ldots, m$이다.

정리 3.4 관측벡터의 조건부결합분포가 식 (3.6)과 같이 주어지는 GBM 모형에서, 확률효과 $\eta^{(k)}$의 분포가 식 (3.7)과 같이 주어질 때, 관측벡터 $\boldsymbol{y}^{(k)}$에 대한 주변분포의 밀도함수 h는 다음과 같다.

$$h\left(\boldsymbol{y}^{(k)}|\beta, \nu, \tau\right) = \frac{2}{(2\pi)^{\frac{n}{2}}\Gamma(\nu/2)(2\delta)^{\frac{n}{2}}(2\tau)^{\frac{\nu}{2}}} \exp\left\{-\frac{n}{2}\bar{b}^{(k)}\right\} \cdot \left(\frac{\delta}{B^{(k)}}\right)^{\frac{1}{2}(n+\nu)} K_{\frac{1}{2}(n+\nu)}\left(\frac{n}{2}B^{(k)}\right),$$

여기서 $K_r(\cdot)$는 수정된 베셀함수(MODIFIED BESSEL FUNCTION)이고,

$$B^{(k)} = \left(\overline{b^2}^{(k)} + \frac{2\delta}{n\tau}\right)^{\frac{1}{2}}$$

이고, $\bar{b}^{(k)} = (1/n)\sum_{i=1}^n b_i^{(k)}$이고 $\overline{b^2}^{(k)} = (1/n)\sum_{i=1}^n (b_i^{(k)})^2$이다.

증명: 전체 증명의 내용은 생략한다. 다만 적분과정에서 다음과 같은 수정된 베셀함수의 성질이 이용되었다.

$$\int_0^\infty z^{r-1}\exp\left\{-\frac{1}{2}\left(pz + qz^{-1}\right)\right\}dz = 2\left(\frac{q}{p}\right)^{\frac{r}{2}} K_r\left(\sqrt{pq}\right)$$

여기서, $p, q > 0$이고, r은 임의의 실수이고, $K_{-r}(x) = K_r(x)$이다. □

4. 수치 예제

4.1. 모의실험

위 결과를 확인하기 위하여, 모의실험과 실제 자료에 대한 분석을 하였다. OU 모형, GBM 모형 각각에 대하여, 평균모수(beta) 확률효과모형, 확산모수(eta) 확률효과모형을 모의실험하였다. 이를 기호로 각각 OU(beta), OU(eta), GBM(beta), GBM(eta)로 나타내기로 한다.

모의실험의 각 경우에서 모수의 참값이 다음과 같은 경우를 실험하였다. 모수의 값은 Ait-Sahalia (2002), Lee 등 (2014) 등의 관련 연구에 사용된 경우와 비슷하게 값을 설정하고, 분산을 의미하는 η와 τ 값에 대한 제곱근이 간단하게 표현되는 경우로 변경하였다.

$$\begin{aligned}
\text{OU(beta)} : \theta &= (\alpha, \beta, \eta, \tau) = (0.2, 0.5, 0.09, 0.01),\\
\text{OU(eta)} : \theta &= (\alpha, \beta, \nu, \tau) = (0.2, 0.5, 7.0, 2.2),\\
\text{GBM(beta)} : \theta &= (\beta, \eta, \tau) = (0.8, 0.04, 0.01),\\
\text{GBM(eta)} : \theta &= (\beta, \nu, \tau) = (0.8, 7.0, 5.0).
\end{aligned}$$

확산과정에 대한 관측간격 δ는 0.01로 가정하였다. 표본의 크기, (m, n)은 $(30, 50)$, $(50, 50)$, $(30, 100)$, $(50, 100)$ 네 가지 경우를 모의시험 하였다. 여기서 n은 관측벡터의 크기이고, m은 반복하여 관측된 관측벡터의 개수이다.

모형에 사용된 모수 α, η, τ, ν은 양수라는 제약 조건을 갖는 값으로, 모수의 추정을 위하여, 범위 제약이 있는 경우에서 함수최적화를 달성하는 R의 nlminb 함수를 이용하여 모수를 추정하였다. 모의실험에서의 반복횟수는 $l = 1000$이다.

각 모의실험 결과로 얻은 모수의 추정값 , 들에 대하여 평균과 평균제곱오차 제곱근을 표로 정리하였다. 표에 나타난 평균(mean)과 평균제곱오차 제곱근(root mean squared error; RMSE)은 다음과 같이 구하였다.

$$\text{MEAN} = \frac{1}{l}\sum_{s=1}^{l}\hat{\theta}_s, \qquad \text{RMSE} = \sqrt{\frac{1}{l}\sum_{s=1}^{l}\left(\hat{\theta}_s - \theta\right)^2}.$$

Table 4.1은 OU 모형의 경우, Table 4.2는 GBM 모형의 경우이다. 표본의 크기 n과 m이 커질수록 MEAN이 참값과 가까워지고 있고, RMSE 값이 줄어드는 경향을 보이고 있다. 확률효과의 모수인 경우나 그렇지 않은 경우 모두 표본의 크기 n과 m에 영향을 받아 RMSE가 줄어드는 모습을 보이고 있다.

4.2. 다우존스 산업지수의 예

1991년 1월부터 2017년 12월까지의 27년 324개월 동안의 다우존스지수(DJIA)를 이용하였다. 매일 매일의 종가 자료를 이용하여 GBM 모형을 적용하였다. 관측벡터의 크기 n은 매년 뉴욕증권거래소의 개장일 수를 의미하고, 관측벡터의 개수 m은 27이다. 매일 매일 관측된 확산과정이므로, 관측과정을 1년을 시간의 단위로 하여, 일일 자료의 관측간격 δ를 1/252로 계산하였다.

다우존스지수 자료에 대하여, GBM(beta) 모형과 GBM(eta) 모형을 적용하여 모수를 추정한 결과는 다음과 같다.

$$\begin{aligned}
\text{GBM(beta)} : \theta &= (\beta, \eta, \tau) = (0.0942, 0.0141, 0.0000),\\
\text{GBM(eta)} : \theta &= (\beta, \nu, \tau) = (0.1322, 4.0168, 32.1064).
\end{aligned}$$

Table 4.1. Simulation results for OU models

n	m	OU(beta)					OU(eta)			
		θ	α	β	η	τ	α	β	η	τ
		True	0.2	0.5	0.09	0.01	0.2	0.5	7.0	2.2
50	30	MEAN	0.2524	0.4742	0.1149	0.2897	0.2477	0.6128	7.9797	2.5936
		RMSE	0.1550	0.4318	0.1162	1.2801	0.1662	7.2895	2.6405	1.9388
50	50	MEAN	0.2201	0.4991	0.1211	0.2477	0.2202	0.3637	7.5906	2.3461
		RMSE	0.1151	0.3365	0.1146	1.2455	0.1252	2.9571	1.8835	1.4725
100	30	MEAN	0.2257	0.5074	0.1186	0.1917	0.2262	0.5381	7.9265	2.3862
		RMSE	0.1207	0.3251	0.1152	1.4923	0.1232	3.5769	2.6467	1.5129
100	50	MEAN	0.2063	0.4982	0.1124	0.0943	0.2118	0.6335	7.4797	2.3059
		RMSE	0.0889	0.2096	0.0825	0.4375	0.0901	3.1485	1.7386	1.1641

OU = Ornstein-Uhlenbeck; RMSE = root mean squared error.

Table 4.2. Simulation results for GBM models

n	m	GBM(beta)				GBM(eta)		
		θ	β	η	τ	β	ν	τ
		True	0.8	0.04	0.01	0.8	7.0	5.0
50	30	MEAN	0.7995	0.0399	0.0194	0.7992	7.9188	4.8844
		RMSE	0.0761	0.0013	0.0305	0.0641	2.9334	1.5981
50	50	MEAN	0.7968	0.0399	0.0163	0.7974	7.5254	4.8699
		RMSE	0.0575	0.0011	0.0249	0.0494	1.7613	1.1566
100	30	MEAN	0.7971	0.0399	0.0138	0.7995	7.8843	4.8043
		RMSE	0.0555	0.0011	0.0178	0.0427	2.5306	1.4097
100	50	MEAN	0.7991	0.0399	0.0121	0.7991	7.5302	4.9055
		RMSE	0.0414	0.0008	0.0141	0.0329	1.8415	1.1603

GBM = geometric Brownian motion; RMSE = root mean squared error.

평균모수를 확률효과로 가정하는 경우, $\beta^{(k)}$의 분산, τ가 0으로 추정되어, $\beta^{(k)}$는 확률효과를 갖지 않는 퇴화된 경우로 추정되었고, 확산모수를 확률효과로 가정하는 경우, 확률효과 $\eta^{(k)}$는 그 역수의 분포가 자유도가 4.02인 카이제곱분포에, 규모모수가 32.11인 것으로 추정되었다. 즉, $(\eta^{(k)})^{-1}$는 평균이 128.97이고, 표준편차 64.35인 것으로, $\eta^{(k)}$의 평균은 0.0154, 표준편차 0.1525인 것으로 추정되었다.

확률효과를 가정한 평균모수의 분산이 0으로 추정된 것은, 다우존스 자료로 대표되는 주가지수 자료의 경우 평균모수에 대한 확률효과는 없는 것으로 볼 수 있음을 의미한다. 확산모수 확률효과 모형의 경우, 평균에 비하여 표준편차가 매우 크게 나타나고 있는 것은, 대부분의 시기에 있어서 확산모수의 값이 일정한 값을 유지하고 있으나, 금융위기 상황과 같은 특정시기에 있어서 주가의 변동성이 매우 높게 나타나는 경우가 있기 때문인 것으로 생각된다.

5. 맺음말

본 논문에서는, 가장 기본적인 확산모형인 OU 모형, GBM 모형에 대한 확률효과 모형을 살펴보았다. 또한 모형모수들이 동시에 확률효과를 가질 수 있다고 보는 대신, 개별 모수 각각이 확률효과가 되는 경우만 살펴보았다. 그런 점에서 본 논문에서의 연구는 일반성이 부족하고 매우 제약적이라고 할 수 있다.

그러나, OU 모형과 GBM 모형이 가장 기본적인 확산모형이고, 본 논문에서 주장하는 바의 취지는, 확산현상을 모형화하기 위하여 복잡한 확산과정을 고려하는 대신, 단순한 확산과정을 이용하고, 모형모수

에 확률효과를 도입하는 것이 바람직하다는 것이므로, 본 논문에서 단순한 확산과정만을 고려하였다 하여 논문의 취지가 훼손되는 것은 아니다. 오히려 본 논문에서 제시한 OU 모형과 GBM 모형에 확률효과를 도입한 경우의 우도함수는 실제 확산현상을 설명하거나 확산현상에 대한 자료를 해석하는데 유용하게 자주 사용될 수 있을 것으로 예상된다.

본 논문을 통하여, 매우 단순한 형태의 확산모형에 대하여도, 확률효과를 도입하는 경우, 적분의 형태로 주어지는 관측벡터의 주변분포(즉, 우도함수)를 구하는 것이 쉽지 않음을 살펴볼 수 있다. 이는 일반적인 확산모형에 확률효과를 도입하는 경우 그 모수 추정이 단순하고 쉬운 문제가 아니라는 의미가 되는 것이고, 일반적인 확산모형에 확률효과를 고려하기 위해서는 별도의 통계적 방법이 개발되어야 함을 의미한다.

본 연구에서는 이후에 보다 깊이 있는 연구가 진행되는 과정에서 참고가 되고 결과에 대한 비교 기준이 될 수 있도록, 가장 단순한 확산모형인 OU 모형과 GBM 모형에 대하여 확률효과를 도입하는 경우, 적절한 확률효과의 분포를 가정하여 비교적 단순한 방법으로 그 우도함수의 형태를 수식으로 표현할 수 있는 경우를 살펴보았다.

References

Ait-Sahalia, Y. (2002). Maximum-likelihood estimation of discretely-sampled diffusions: a closed-form approximation approach, *Econometrica*, **70**, 223–262.

Applebaum, D. (2004) Levy processes - from probability theory to finance and quantum groups, *Notices of the American Mathematical Society*, **51**, 1336–1347.

Delattre, M., Genon-Catalot, V., and Samson, A. (2013). Maximum likelihood estimation for stochastic differential equations with random effects, *Scandinavian Journal of Statistics*, **40**, 322–343.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options, *Reviews of Financial Studies*, **6**, 327–343.

Hurn, A., Jeisman, J., and Lindsay, K., (2007). Seeing the wood for the trees: a critical evaluation of methods to estimate the parameters of stochastic differential equations, *Journal of Financial Econometrics*, **5**, 390–455.

Laird, N. M. and Ware, J. H. (1982). Random-effects models for longitudinal data, *Biometrics*, **38**, 963–974.

Lee, Y. D., Song, S., and Lee, E. (2014). The delta expansion for the transition density of diffusion models, *Journal of Econometrics*, **178**, 694–705.

Picchini, U., De Gaetano, A., and Ditlevsen, S. (2010). Stochastic differential mixed-effects models, *Scandinavian Journal of Statistics*, **37**, 67–90.

Pinheiro, J. C. and Bates, D. M. (2000). *Mixed-Effects Models in S and S-PLUS*, Springer-Verlag, New York.

단순 확산과정들에 대한 확률효과 모형

이은경[a] · 이인석[b] · 이윤동[b,1]

[a]이화여자대학교 통계학과, [b]서강대학교 경영학부

(2018년 10월 29일 접수, 2018년 11월 7일 수정, 2018년 11월 7일 채택)

요 약

확산은 금융이나 물리적 현상의 모형화에 이용되는 확률과정이다. 반복적으로 관측된 확산과정에 대하여 통계적인 모형을 구축할 때, 확률효과를 고려할 필요가 있다. 이 연구에서는 Ornstein-Uhlenbeck 확산모형과 geometric Brownian motion 확산모형에 대하여 확률효과를 도입한다. 모형모수에 대한 최도우도추정법을 적용하기 위하여, 확률효과에 대한 적절한 분포를 가정하여 닫힌 형태로 우도함수를 얻는 방법을 탐색하였다. 1991년부터 2017년까지 27년간 일일 단위로 기록된 다우존스 산업지수에 대하여 확률효과 모형을 적용하였다.

주요용어: 확산, 확률효과, OU 모형, GBM 모형

이 논문은 2015년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (2015R1D1A1A01056790).

[1]교신저자: (04107) 서울시 마포구 백범로 35, 서강대학교 경영학부. E-mail: widylee@sogang.ac.kr