UNIVERSITY OF COPENHAGEN

Some Simple ML Estimators in Stochastic Differential Equations

Andersen, Erling B.

Publication date:
2001

Document version
Early version, also known as pre-print

Citation for published version (APA):
Andersen, E. B. (2001). *Some Simple ML Estimators in Stochastic Differential Equations*. Department of Economics, University of Copenhagen.

DISCUSSION PAPERS
Department of Economics
University of Copenhagen

01-10

Some Simple ML Estimators in Stochastic Differential Equations

Erling B. Andersen

Studiestræde 6, DK-1455 Copenhagen K., Denmark
Tel. +45 35 32 30 82 - Fax +45 35 32 30 00
http://www.econ.ku.dk

Some simple ML-estimators in stochastic differential equations.

by Erling B. Andersen
University of Copenhagen

Abstract: For many stochastic differential equations, often met in financial theory, its is the drift and the dispersion which are the principal parameters of the model. In such cases it is shown that the parameters can be estimated by ordinary methods from normal distribution theory.

Key words: Stochastic differential equaltions. ML-estimates. Financial models.

1. Introduction

A general stochastic differential equation for a stochastic process W(t) is defined as

$$dW(t) = f\big(t, W(t)\big)dt + g\big(t, W(t)\big)dB(t) \ ,$$

where f and g are known functions of their arguments and B(t) is a Brownian motion or a Wienerprocess.

There has been considerable interest over the last decade in estimation problems connected with parametrized stochastic differential equations.

The basic approach has been to parametrize the stochastic process itself, i.e. to assume that W(t) depends on a vector of parameters $\theta = (\theta_1, \ldots ,\theta_k)$, such that the stochastic differential equation becomes

$$dW(t|\theta) = f\big(t, W(t|\theta)\big)dt + g\big(t, W(t|\theta)\big)dB(t) \ , \qquad (1)$$

For this parametric representation of a stochastic differential equation several estimation procedures has been suggested. From an applied point of view, most of these are rather complicated and involve advanced results from the theory of diffusion processes. Cf. e.g. Kessler and Sørensen (1999) for important results and references.

In this paper, we consider a subclass of stochastic differential equations, which, although simple in structure, covers several interesting cases from applications to financial theory.

Consider thus the alternative formulations to (1)

$$dW(t) = \alpha f(t)dt + \sigma g(t)dB(t) \ , \tag{2}$$

where f(t) and g(t) are known functions of t, or

$$dW(t) = \alpha f\big(t,W(t)\big)dt + \sigma g\big(t,W(t)\big)dB(t) \ , \tag{3}$$

where again f and g are known functions of their arguments.

The purpose of this paper is to demonstrate, that for simple choices of the functions f and g in (2) and (3), it is possible to derive simple estimators for α and σ. In addition the distributional properties of these estimators are easily derived from ordinary normal distribution theory.

For simplicity, we shall mainly consider the equidistant case, where the process is observed at times t = 1, t = 2, t = 3 and so on. The results are easily extended to cases with non-equidistant observation points in time. For the case in Section 2 it is demonstra-ted, how this can be done.

We shall call the parameter α in (2) or (3) the drift and the parameter σ in (2) or (3) the dispersion of the process.

2. Constant drift and constant dispersion

If f(t) = 1 and g(t) = 1, we shall say that the process has constant drift and constant dispersion. In this case Equation (2) takes the form

$$dW(t) = \alpha dt + \sigma dB(t) \ , \tag{4}$$

or

$$\int_0^t W(s)\,ds \;=\; \alpha \int_0^t ds \;+\; \sigma \int_0^t dB(t) \ .$$

The last integral is an Ito-integral, but in this case the function to be integrated is constant, such that the solution is B(t) - B(0). Hence the solution to (4) is

$$W(t) \;=\; W(0) \;+\; \alpha t \;+\; \sigma\big(B(t) - B(0)\big) \;=\; W(0) \;+\; \alpha t \;+\; \sigma B(t),$$

since B(0) = 0, by the definition of a Brownian motion.

Suppose now, that W(t) has been observed at times t = 1, 2, ... , n, and that the observed values are w(1), w(2), ... , w(n). With w(0) = 0, we can then form the new observations y_1, ... ,y_n as

$$y_i \;=\; w(i) \;-\; w(i-1) \;=\; \alpha(i-(i-1)) \;+\; \sigma\big(b(i) - b(i-1)\big) \ ,$$

or

$$y_i \;=\; \alpha \;+\; \sigma\big(b(i) - b(i-1)\big) \quad , \quad i = 1,\ldots,n \ . \tag{5}$$

where b(i) is the realized value of the Brownian motion B(i) at time i.

The random variables Y_1,...,Y_n corresponding to the observed y’s thus have the common distribution

$$Y_i \sim N\big(\alpha,\sigma^2\big) \ . \tag{6}$$

The variance of Y_i follows from the fact, that for a Brownian motion the variance at time t is var[B(t)]=t, and the covariance between time t and time s, t < s, is

$$\mathrm{cov}\big(B(t),B(s)\big) \;=\; t \ .$$

Accordingly

$$\mathrm{var}\left[B(i) - B(i-1)\right] = i + (i-1) - 2(i-1) = i - (i-1) = 1 .$$

In addition the Y's are independent, since Y_i = W(i) - W(i-1) = $\alpha + \sigma$(B(i) - B(i-1)) and a Brownian motion has independent increments.

If we write w(i) = w_i, the transformation from (y_1, ... ,y_n) to (w_1, ... ,w_n) is one-one and the differential element is

$$\frac{dy_1 \ldots dy_n}{dw_1 \ldots dw_n} = 1.$$

Hence with the proper substitutions

$$f_Y\left(y_1, \ldots, y_n\right) = f_W\left(w_1, \ldots, w_n\right) .$$

This means that the likelihood function for the y's is equal to the likelihood function for the w's. The estimation of α and σ and the distridutional properties of these parameters can thus be derived from the distribution of Y_1, ... ,Y_n, which is a set of independent, identically and normally distributed random variables.

The ML-estimator of α is thus

$$\hat{\alpha} = \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{w(n) - w(0)}{n} = \frac{w(n)}{n} , \tag{7}$$

since w(0) = 0. The ML-estimator of σ^2 is

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} \left(y_i - \bar{y}\right)^2 = \frac{1}{n}\sum_{i=1}^{n} \left(\triangle w(i) - \frac{w(n)}{n}\right)^2 , \tag{8}$$

where

$$\triangle w(i) = w(i) - w(i-1) .$$

The distribution of the estimators follows from (6). Thus

$$\hat{\alpha} \sim N\left(\alpha, \frac{\sigma^2}{n}\right)$$

and the distribution of $\hat{\sigma}$ is described by

$$\frac{n\hat{\sigma}^2}{\sigma^2} = \frac{n-1}{n} s^2 \frac{n}{\sigma^2} = \frac{(n-1)s^2}{\sigma^2} ,$$

where s^2 is the unbiased estimator for σ^2. Hence

$$\frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-1) .$$

For non-equidistant time points $t_1, \ldots, t_n$, the observations are $w(t_1), w(t_2), \ldots, w(t_n)$. With $w(t_0) = 0$, we again base the estimations on the differences

$$y_i = w(t_i) - w(t_{i-1}) = \alpha(t_i - t_{i-1}) + \sigma\left(b(t_i) - b(t_{i-1})\right) ,$$

or

$$y_i = \triangle w(t_i) = \alpha \cdot \triangle t_i + \sigma \cdot \triangle b(t_i) , \tag{9}$$

where

$$\triangle w(t_i) = w(t_i) - w(t_{i-1}) \quad , \quad \triangle t_i = t_i - t_{i-1} \quad , \quad \triangle b(t_i) = b(t_i) - b(t_{i-1}) .$$

From (9) follows that

$$E\left[Y_i\right] = \alpha \cdot \triangle t_i$$

and

$$\mathrm{var}\left[Y_i\right] = \sigma^2\left(t_i + t_{i-1} - 2t_{i-1}\right) = \sigma^2\left(t_i - t_{i-1}\right) = \sigma^2 \triangle t_i ,$$

since

$$\mathrm{cov}\left(B(t_i), B(t_{i-1})\right) = t_{i-1} .$$

In addition the Y’s are independent since a Brownian motion has independent increments.

The log-likelihoodfunction is accordingly

$$\ln L = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2}\sum_{i=1}^{n} \ln\left(\triangle t i\right) - \frac{1}{2}\sum_{i=1}^{n} \frac{\left(\triangle w(t_i) - \alpha \triangle t_i\right)^2}{\sigma^2 \triangle t_i}$$

The ML-estimates for α and σ^2 are easily derived from ln L as

$$\hat{\alpha} = \frac{\sum_i \triangle w(t_i)}{\sum_i \triangle t_i} = \frac{w(t_n) - w_0}{t_n - 0} = \frac{w(t_n - w_0}{t_n} \tag{10}$$

and

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} \frac{\left(\triangle w(t_i) - \hat{\alpha} \triangle t_i\right)^2}{\triangle t_i} \;. \tag{11}$$

Since $w(t_0) = 0$

$$\hat{\alpha} = \frac{w(t_n)}{t_n} \;, \tag{12}$$

and inserting (12) in (11)

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{\triangle t_i}\left(\triangle w(t_i) - \frac{w(t_n)}{t_n}\triangle t_i\right)^2 \;.$$

In order to derive the distribution of $\hat{\sigma}$ we rewrite (11) as

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\triangle w(t_i)}{\sqrt{\triangle t_i}} - \hat{\alpha}\sqrt{\triangle t_i}\right)^2 \;.$$

Here

$$\operatorname{var}\left[\frac{\triangle W(t_i)}{\sqrt{\triangle t_i}}\right] = \frac{1}{\triangle t_i}\triangle t_i \sigma^2 = \sigma^2 ,$$

while

$$E\left[\frac{\triangle W(t_i)}{\sqrt{\triangle t_i}}\right] = \frac{1}{\sqrt{\triangle t_i}}\triangle t_i \alpha = \sqrt{\triangle t_i}\,\alpha ,$$

This means that we can perform the classical splitting of sum of squares, necessary to obtain the distribution of the unbaised estimate for σ^2. Her we get

$$Q = \sum_{i=1}^{n}\left(\frac{\triangle w(t_i)}{\sqrt{\triangle t_i}} - \alpha\sqrt{\triangle t_i}\right)^2$$

$$= \sum_{i=1}^{n}\left(\frac{\triangle w(t_i)}{\sqrt{\triangle t_i}} - \hat{\alpha}\sqrt{\triangle t_i} + \hat{\alpha}\sqrt{\triangle t_i} - \alpha\sqrt{\triangle t_i}\right)^2$$

$$= \sum_{i=1}^{n}\left(\frac{\triangle w(t_i)}{\sqrt{\triangle t_i}} - \hat{\alpha}\sqrt{\triangle t_i}\right)^2 + (\hat{\alpha} - \alpha)^2 t_n = Q_1 + Q_2 .$$

By the addition theorem for the χ^2-distribution

$$\frac{Q}{\sigma^2} \sim \chi^2(n) ,$$

while obviously

$$\frac{Q_2}{\sigma^2} \sim \chi^2(1) .$$

Hence by Cochrans theorem

$$\frac{Q_1}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n}\left(\frac{\triangle w(t_i)}{\sqrt{\triangle t_i}} - \hat{\alpha}\sqrt{\triangle t_i}\right)^2 \sim \chi^2(n-1).$$

or

$$\frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-1) \ .$$

The results for non-equidistant observed time points are thus similar to those for equi-distantly observed time points.

3. The drift a function of t and constant dispersion

For these cases, we assume, that the drift is an integrable function f(t) of t, and that the dispersion is constant, i.e. the stochastic differential equation is given by

$$dW(t) = \alpha f(t)dt + \sigma dB(t) \ .$$

By integration we then get

$$W(t) = W(0) + \alpha\int_0^t f(s)ds + \sigma\big(B(t) - B(0)\big)$$

i.e. with F(t) being the integral

$$F(t) = \int_0^t f(s)ds \ ,$$

and since B(0) = 0

$$W(t) = W(0) + \alpha F(t) + \sigma B(t) \ .$$

Taking differences, we get with t = i = 1, 2, 3, ...

$$Y_i = W(i) - W(i-1) = \triangle W(i) = \alpha \triangle F(i) + \sigma \triangle B(i) ,$$

where as before $\Delta F(i) = F(i) - F(i-1)$. The Y's are independent, since the B-differences are independent. With the $z_i = \Delta F(i)$ being known 'explanatory variables and

$$\triangle B(i) \sim N(0,1) ,$$

the model is thus equivalent to an origo regression model with explanatory variables $z_1, \ldots , z_n$, and the estimates and the distributional properties follow from ordinary regression analysis theory. This means that the ML-estimators are given by

$$\hat{\alpha} = \frac{\sum_i z_i y(i)}{\sum_i z_i^2} ,$$

and

$$\hat{\sigma}^2 = \frac{1}{n}\left(\sum_i y_i^2 - \frac{\sum_i z_i y_i}{\sum_i z_i^2} \right) .$$

For the special case $f(t) = t$, we get

$$F(t) = \int_{t-1}^{t} s\,ds = \frac{t^2}{2} - \frac{(t-1)^2}{2} = t - \frac{1}{2} ,$$

or the midpoint of the interval from t-1 to t.

4. The drift and the dispersion proportional to t

In this case, we consider the stochastic differential equation

$$dX(t) = \mu t dt + \sigma t dZ(t) .$$

Here it pays to study directly the differences

$$\int_{t-1}^{t} dW(t) = W(t) - W(t-1) = \triangle W(t) = \alpha \int_{t-1}^{t} s\,ds + \sigma \int_{t-1}^{t} s\,dB(t)$$

$$= \alpha\left(t - \frac{1}{2}\right) + \sigma \int_{t-1}^{t} s\,dB(t)$$

From known properties of Brownian motions, cf. Øksendahl (1998), Section 3.2, follows

$$E\left[\int_{t-1}^{t} s\,dZ(s)\right] = 0 \ . \tag{13}$$

Hence by Iso's isometry

$$\mathrm{var}\left[\int_{t-1}^{t} s\,dZ(s)\right] = E\left[\left(\int_{t-1}^{t} s\,dZ(s)\right)^2\right] = E\left[\int_{t-1}^{t} s^2 ds\right]$$

$$= \int_{t-1}^{t} s^2 ds = \frac{1}{3}\left(t^3 - (t-1)^3\right) = t(t-1) - \frac{1}{3} \ .$$

Accordingly the independent observations y_i = w(i) - w(i-1) have distributions

$$Y_i \sim N\left(\alpha z_i, c_i \sigma^2\right) \ ,$$

with z_i = i - 1/2 and c_i = i(i-1) - 1/2.

The model is thus equivalent to a weighted origo-regression model with weights $1/c_i$ and 'explanatory variables' z_i. The ML-estimates for α and σ^2 are, therefore,

$$\hat{\alpha} = \frac{\sum_i \frac{1}{c_i} y_i z_i}{\sum_i \frac{1}{c_i} z_i^2}$$

and

$$\hat{\sigma}^2 = \frac{1}{n}\left(\sum_i \frac{1}{c_i} y_i^2 - \frac{\sum_i \frac{1}{c_i} z_i y_i}{\sum_i \frac{1}{c_i} z_i^2}\right)$$

5. The drift and the dispersion proportional to W(t)

Consider finally the model

$$dW(t) = \alpha W(t)dt + \sigma W(t)dB(t) . \tag{14}$$

where both the drift and the dispersion are proportional to the process itself. This model can also be written as

$$\frac{dW(t)}{W(t)} = \alpha dt + \sigma dB(t) ,$$

In order to evaluate the left hand side, we must apply Ito’s lemma to the function lnW(t), cf. Øksendahl (1998), Section 5.1, Example 5.1.1. From Ito’s lemma, we get

$$d\left(\ln W(t)\right) = \frac{dW(t)}{W(t)} + \frac{1}{2}\left(-\frac{1}{W(t)^2}\right)\left(dW(t)\right)^2 \tag{15}$$

But according to Ito’s identities and (14)

$$\left(dW(t)\right)^2 = \alpha^2 W(t)^2 (dt)^2 + \sigma^2 W(t)^2 \left(dB(t)\right)^2 +$$
$$2\alpha W(t)\sigma dt dB(t) = \sigma^2 W(t)^2 dt .$$

Thus from (15)

$$d\left(\ln W(t)\right) = \frac{dW(t)}{W(t)} - \frac{1}{2}\sigma^2 dt = \left(\alpha - \frac{1}{2}\sigma^2\right)dt + \sigma dB(t) .$$

If we, therefore, use the observed values

$$y_i = \ln w(i) - \ln w(i-1)$$

of the logarithmically transformered variables as observations, with the new trendparameter

$$\alpha_1 = \alpha - \frac{1}{2}\sigma^2 .$$

we get the ML-estimates

$$\hat{\alpha}_1 = \bar{y} ,$$

and

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i}^{n}\left(y_i - \bar{y}\right)^2 .$$

for α_1, and for σ^2 .

The estimate for the origional α becomes

$$\hat{\alpha} = \hat{\alpha}_1 + \frac{1}{2}\hat{\sigma}^2 .$$

6. References

Kessler, M. and Sørensen, M. (1999): Estimating equations based on eigenfunctions for a discretely observed diffusion process. *Bernoulli.* **5**, 299-314.

Øksendahl, B. (1998): *Stochastic Differential Equations. An Introduction with Applications.* Fifth Edition. Berlin: Springer Verlag.