SUMMER SCHOOL

Stochastic Differential Equation Models

Middelfart 4-12 August 2008

Lectures by

Martin Jacobsen

University of Copenhagen

HOMOGENEOUS DIFFUSIONS

Let $\mathbf{B} = (B_t)_{t\geq 0}$ be an n-dimensional Brownian motion on $(\Omega, \boldsymbol{F}, \boldsymbol{F}_t, P)$. Consider the SDE

$$\underbrace{dX_t}^{n\times 1} = \underbrace{b(X_t)}^{n\times 1}\,dt + \underbrace{\sigma(X_t)}^{n\times n}\underbrace{dB_t}^{n\times 1} \tag{1}$$

with $b : D \to \mathbf{R}^n$ and $\sigma : D \to \mathbf{R}^{n\times n}$ continuous functions defined on an open subset $D \subseteq \mathbf{R}^n$ with $\sigma(x)$ non-singular for all $x \in D.b$ is the *drift* function, σ the *diffusion* function.

Assume that for every $x_0 \in D$, (1) has a unique strong solution with $X_0 \equiv x_0$ such that

$$P^{x_0}\bigcap_{t\geq 0}(X_t \in D) = 1.$$

$\mathbf{X}$ is an n-dimensional *homogeneous diffusion* with *state space* D (an ill-defined concept). If $X_0 \equiv x_0$, write P^{x_0} instead of P; if X_0 is random with distribution ν, write P^{ν} instead of P.

Existence and uniqueness of solution process $\mathbf{X}$: OK if $D = \mathbf{R}^n$ and b and σ are Lipschitz and of linear growth. Important problem: for general D, what are the conditions that must be imposed on b and σ in order for $\mathbf{X}$ to stay inside D and never hit the boundary of D? We shall answer this for $n = 1$ with $D = \,]l, r[$, allowing also for $l = -\infty$ or $r = \infty$ — for $n \geq 2$ there are no general results.

Note. It is perfectly possible to define e.g one-dimensional diffusions with a state interval that is closed or half-open ($[l, r]$ or $]l, r]$, $[l, r[$), diffusions with reflecting or absorbing boundaries. Here we focus on open intervals $]l, r[$.

Let Q^{x_0} denote the *distribution* of $\mathbf{X}$ when $X_0 \equiv x_0$, ie Q^{x_0} is the probability measure on $\left(W^{(n)}, \boldsymbol{H}^{(n)}\right)$ given by

$$Q^{x_0}(H) = P^{x_0}(\mathbf{X} \in H) \quad \left(H \in \boldsymbol{H}^{(n)}\right).$$

Here $W^{(n)}$ is the space of all continuous functions $w : [0, \infty[\to \mathbf{R}^n$ and $\boldsymbol{H}^{(n)}$ is the smallest σ-algebra of subsets making all mappings $w \mapsto w(t)$ for $t \geq 0$ measurable.

$\mathbf{X}$ AS A MARKOV PROCESS

$\mathbf{X}$ is a *time-homogeneous Markov process:* for all $t \geq 0$, $H \in \boldsymbol{H}^{(n)}$,

$$P^{x_0}\left((X_{t+u})_{u\geq 0} \in H \,|\boldsymbol{F}_t\right)(\omega) = Q^{X_t(\omega)}(H).$$

$\mathbf{X}$ is even *strong Markov:* for all stopping times τ, if $\tau(\omega) < \infty$,

$$P^{x_0}\left((X_{\tau+u})_{u\geq 0} \in H \,|\boldsymbol{F}_\tau\right)(\omega) = Q^{X_\tau(\omega)}(H).$$

Example. If τ is a finite stopping time, $(B_{\tau+u} - B_\tau)_{u\geq 0}$ is a Brownian motion independent of $\boldsymbol{F}_\tau$.

Define the *transition probabilities*

$$p_s(x, B) = P^x(X_s \in B) \quad (B \in \mathbf{B}^n).$$

Then (take $H = \left\{w \in W^{(n)} : w(s) \in B\right\}$)

$$P^x(X_{t+s} \in B \,|\boldsymbol{F}_t)(\omega) = p_s(X_t(\omega), B)$$

and

$$\begin{aligned} p_{t+s}(x, B) &= E^x p_s(X_t, B) \\ &= \int_D p_s(y, B)\, p_t(x, dy), \end{aligned}$$

the *Chapman-Kolmogorov equations.*

The *transition operators* act on bounded $\mathbf{R}$-valued Borel functions,

$$P_t f(x) = E^x f(X_t) = \int_D f(y)\, p_t(x, dy)$$

and form a *semigroup* (the Chapman-Kolmogorov equations),

$$P_{t+s} f \equiv P_t(P_s f) \equiv P_s(P_t f).$$

The *infinitesimal generator* A acts on bounded f such that the following limit exists for all $x \in D$,

$$Af(x) = \lim_{t\downarrow 0} \frac{1}{t}(P_t f(x) - f(x)).$$

Assume that $f \in C^2$ (twice differentiable with continuous second order partial derivatives). By Itô's formula, if $X_0 \equiv x_0$,

$$f(X_t) = f(x_0) + \int_0^t \widetilde{A} f(X_s) + M_t$$

where

$$\widetilde{A} f(x) = \sum_{i=1}^n b_i(x) \frac{\partial}{\partial x_i} f(x) + \frac{1}{2} \sum_{i,j=1}^n c_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} f(x)$$

where $C = \left(c_{ij}\right) = \sigma\sigma^\top$ and $\mathbf{M}$ is the local martingale

$$M_t = \sum_{i,j=1}^n \int_0^t \frac{\partial}{\partial x_i} f(X_s) \sigma_{ij}(X_s)\, dB_{j,s}.$$

If both f and $\widetilde{A} f$ are bounded on D, $\mathbf{M}$ is a true martingale and

$$P_t f(x_0) = f(x_0) + \int_0^t P_s\left(\widetilde{A} f\right)(x_0)\, ds \qquad (2)$$

which implies that

$$\lim_{t\downarrow 0} \frac{1}{t}(P_t f(x_0) - f(x_0)) = \widetilde{A} f(x_0).$$

For such f therefore

$$Af \equiv \widetilde{A} f.$$

A famous generalisation of (2) is *Dynkin's formula:* if τ is a stopping time with $E^{x_0}\tau < \infty$, then for f as above

$$E^{x_0} f(X_\tau) = f(x_0) + E^{x_0} \int_0^\tau Af(X_s)\, ds.$$

ONE-DIMENSIONAL DIFFUSIONS

From now on, $n = 1$ and $D = \left]l, r\right[$ is an open interval. Recall that $x \mapsto \sigma(x) \neq 0$ and is continuous, we may therefore assume that $\sigma(x) > 0$ for all $x \in \left]l, r\right[$.

First we look for $f = S \in C^2$ such that $S(\mathbf{X})$ is a local martingale, ie we require that $AS \equiv 0$ on $\left]l, r\right[$,

$$b(x)S'(x) + \frac{1}{2}\sigma^2(x)S''(x) = 0 \quad (l < x < r).$$

If S is strictly increasing this is equivalent to

$$\left(\log S'\right)'(x) = -\frac{2b(x)}{\sigma^2(x)}.$$

Thus

$$S'(x) = \exp\left(-\int_{x_1}^{x} \frac{2b(y)}{\sigma^2(y)}\, dy\right)$$

with $x_1 \in \left]l, r\right[$ an arbitrarily chosen reference point. Any primitive S of S' is called a *scale function* for $\mathbf{X}$. S is not uniquely determined, but if S is a scale function, all others are of the form $\alpha + \beta S$ with $\alpha \in \mathbf{R}$, $\beta > 0$.

Note. Any scale function S is strictly increasing and finite on $\left]l, r\right[$. One can therefore define the limits

$$S(r) = \lim_{x \uparrow r} S(x), \quad S(l) = \lim_{x \downarrow l} S(x)$$

with $-\infty < S(r) \leq \infty$, $-\infty \leq S(l) < \infty$.

Example. For Brownian motion $\mathbf{B}$ itself, $S(x) = x$ is a scale function.

Assume that $X_0 \equiv x_0 \in \left]l, r\right[$. For $c \in \left]l, r\right[$, $l < a < x_0 < b < r$ define

$$\begin{aligned} \tau_c &= \inf\{t \geq 0 : X_t = c\} \\ \tau_{ab} &= \inf\{t \geq 0 : X_t = a \text{ or } X_t = b\} \end{aligned}$$

with $\inf \varnothing = \infty$.

Fact. $P^{x_0}(\tau_{ab} < \infty) = 1$ (even $E^{x_0}\tau_{ab} < \infty$, see below).

The process $\left(S\left(X_{\tau_{ab} \wedge t}\right)\right)_{t \geq 0}$ is a bounded local martingale, hence a true martingale, so

$$E^{x_0} S\left(X_{\tau_{ab} \wedge t}\right) = S(x_0)$$

for $t \geq 0$. Let $t \to \infty$ and use dominated convergence and the fact above to obtain

$$E^{x_0} S\left(X_{\tau_{ab}}\right) = S(x_0).$$

But, again using the fact,

$$E^{x_0} S\left(X_{\tau_{ab}}\right) = S(b)P^{x_0}(\tau_b < \tau_a) + S(a)P^{x_0}(\tau_a < \tau_b)$$

so

$$\begin{aligned} P^{x_0}(\tau_b < \tau_a) &= 1 - P^{x_0}(\tau_a < \tau_b) \\ &= \frac{S(x) - S(a)}{S(b) - S(a)}. \end{aligned}$$

Example. For $\mathbf{B}$ itself,

$$P^{x_0}(\tau_b < \tau_a) = 1 - P^{x_0}(\tau_a < \tau_b) = \frac{x-a}{b-a}.$$

Let S be a given scale function for $\mathbf{X}$. Define the matching *speed measure* m for $\mathbf{X}$ by

$$m(B) = \int_B k(x)\, dx \quad (B \subseteq]l, r[)$$

where

$$k(x) = \frac{2}{\sigma^2(x) S'(x)}.$$

Note. If S is replaced by $\alpha + \beta S$ (with $\beta > 0$), k is replaced $\frac{1}{\beta} k$.

Example. For $\mathbf{B}$ itself, m is 2 (!) times Lebesgue measure.

Fact. For $l < a < b < r$,

$$\begin{aligned} f(x) &= \int_a^x \frac{(Sy - Sa)(Sb - Sx)}{Sb - Sa} k(y)\, dy \\ &\quad + \int_x^b \frac{(Sx - Sa)(Sb - Sy)}{Sb - Sa} k(y)\, dy \end{aligned}$$

is the unique solution for $x \in [a, b]$ of the second order differential equation

$$Af \equiv -1$$

with the boundary conditions $f(a) = f(b) = 0$.

Therefore, for $x_0 \in [a, b]$

$$E^{x_0} f(X_{\tau_{ab}}) = f(x_0) - E^{x_0} \tau_{ab}.$$

But

$$\begin{aligned} E^{x_0} f(X_{\tau_{ab}}) &= f(b) P^{x_0}(\tau_b < \tau_a) + f(a) P^{x_0}(\tau_a < \tau_b) \\ &= 0 \end{aligned}$$

so

$$E^{x_0} \tau_{ab} = f(x_0).$$

Example. For $\mathbf{B}$ we find

$$f(x_0) = (x_0 - a)(b - x_0).$$

Note. From S' and k one can uniquely recover σ^2 and b. Thus S' and k characterise $\mathbf{X}$.

MAIN THEOREM. Let b, σ be continuous functions on $]l, r[$ with $\sigma > 0$, and let S be a scale function with k the matching density for the speed measure. In order that the SDE

$$dX_t = b(X_t)\, dt + \sigma(X_t)\, dB_t, \quad X_0 \equiv x_0$$

for every $x_0 \in]l, r[$ have a unique strong solution such that

$$P^{x_0} \bigcap_{t \geq 0} (X_t \in]l, r[) = 1,$$

it is necessary and sufficient that (i) and (ii) hold:
(i) $S(r) = \infty$ or $S(r) < \infty$, $\int_x^r (S(r) - S(y))\, k(y)\, dy = \infty$;
(ii) $S(l) = -\infty$ or $S(l) > -\infty$, $\int_l^x (S(y) - S(l))\, k(y)\, dy = \infty$.
In both (i) and (ii), $x \in]l, r[$ may be chosen arbitrarily.

THEOREM. Assume that (i) and (ii) from the *MAIN THEOREM* hold.
(a) For any $l < a < x < b < r$,

$$\begin{aligned} P^x(\tau_b < \infty) &= \frac{S(x) - S(l)}{S(b) - S(l)}, \\ P^x(\tau_a < \infty) &= \frac{S(r) - S(x)}{S(r) - S(a)}. \end{aligned}$$

In particular, $P^x(\tau_b < \infty) = 1$ iff $S(l) = -\infty$ and $P^x(\tau_a < \infty) = 1$ iff $S(r) = \infty$.
(b) If $S(r) < \infty$, $S(l) = -\infty$, then for all $x \in]l, r[$

$$P^x\left(\lim_{t\to\infty} X_t = r\right) = 1.$$

If $S(r) = \infty$, $S(l) > -\infty$, then for all $x \in]l, r[$

$$P^x\left(\lim_{t\to\infty} X_t = l\right) = 1.$$

(c) If $-\infty < S(l) < S(r) < \infty$, then for all $x \in]l, r[$

$$\begin{aligned} P^x\left(\lim_{t\to\infty} X_t = r\right) &= 1 - P^x\left(\lim_{t\to\infty} X_t = l\right) \\ &= \frac{S(x) - S(l)}{S(r) - S(l)} \in]0, 1[\,. \end{aligned}$$

(d) If $S(l) = -\infty$, $S(r) = \infty$, $\mathbf{X}$ is *recurrent* and for all $x \in]l, r[$,

$$P^x \bigcap_{c\in]l,r[} \bigcap_{t\geq 0} \bigcup_{s>t} (X_s = c),$$

ie $\mathbf{X}$ passes through all levels $c \in]l, r[$ arbitrarily far out in time.

Example. $\mathbf{B}$ itself is recurrent on $\mathbf{R}$.

Example. The *Ornstein-Uhlenbeck* process is the solution to the SDE

$$dX_t = -\rho X_t\, dt + dB_t$$

where $\rho \neq 0$. This obviously (!) has a unique strong $\mathbf{R}$-valued solution corresponding to any initial condition $X_0 \equiv x_0$ with $x_0 \in \mathbf{R}$.

We find

$$\begin{aligned} S'(x) &= \exp\left(\int_0^x 2\rho y\, dy\right) = e^{\rho x^2}, \\ k(x) &= 2e^{-\rho x^2} \end{aligned}$$

and are forced to take $l = -\infty$, $r = \infty$.

If $\rho > 0$, then $S(l) = -\infty$, $S(r) = \infty$ and $\mathbf{X}$ is recurrent.

If $\rho < 0$, then $-\infty < S(l) < S(r) < \infty$ and

$$\begin{aligned} P^x\left(\lim_{t\to\infty} X_t = \infty\right) &= 1 - P^x\left(\lim_{t\to\infty} X_t = -\infty\right) \\ &= \sqrt{\frac{|\rho|}{\pi}} \int_{-\infty}^x e^{\rho y^2}\, dy. \end{aligned}$$

Exercise. Consider the Ornstein-Uhlenbeck process with $\rho < 0$. According to (i) from the *MAIN THEOREM* it must necessarily hold that

$$\begin{aligned} &\int_x^\infty (S(\infty) - S(y))k(y)\, dy \\ &= \int_x^\infty \int_y^\infty e^{\rho z^2}\, dz\, 2e^{-\rho y^2}\, dy = \infty. \end{aligned}$$

Please verify directly that the double integral is in fact infinite (this is not quite as simple as it sounds)!

Example. For $n \geq 2$, the n-dimensional *Bessel process* is the one-dimensional (!) homogeneous diffusion defined as

$$\mathbf{X} = \|\mathbf{B}\| = \left(\sum_{i=1}^{n} \mathbf{B}_i^2\right)^{\frac{1}{2}}$$

where $\mathbf{B} = (\mathbf{B}_1, \ldots, \mathbf{B}_n)$ is an n-dimensional Brownian motion. It can be shown that $\mathbf{X}$ when staying away from 0 solves the SDE

$$dX_t = \frac{n-1}{2X_t}\,dt + d\widetilde{B}_t$$

with $\widetilde{\mathbf{B}}$ a one-dimensional Brownian motion. This SDE of course fits into the framework of the *MAIN THEOREM* and we (you) now verify that for all integers $n \geq 2$ the SDE defines a homogeneous diffusion on $]l, r[=]0, \infty[$:

$$S'(x) = x^{-(n-1)}, \quad k(x) = 2x^{n-1}$$

and (i), (ii) from the *MAIN THEOREM* follow.

It also follows that for $n = 2$, $\mathbf{X}$ is recurrent while for $n \geq 3$

$$P^x\left(\lim_{t\to\infty} X_t = \infty\right) = 1.$$

For the n-dimensional Brownian motion $\mathbf{B}$ this implies the following: $B_0 = (0, \ldots, 0)$ but immediately after the start, $\mathbf{B}$ leaves the origin and never returns. For $n \geq 3$, $\|B_t\| \to \infty$ as $t \to \infty$ while for $n = 2$, at arbitrarily large times, B_t gets arbitrarily close to the origin without ever getting there.

Fact. If $\mathbf{X}$ is the 2-dimensional Bessel process on $]0, \infty[$ with $X_0 \equiv x_0$, we know that $S(\mathbf{X}) = \log \mathbf{X}$ is a local martingale. It is remarkable however that $\log \mathbf{X}$ is *not* a true martingale: one may show that

$$E^{x_0} \log X_1 = \log x_0 + \int_{x_0}^{\infty} \frac{1}{r} e^{-\frac{1}{2}r^2}\,dr > \log x_0.$$

Exercise. Consider the following generalisation of the SDE for the Bessel processes,

$$dX_t = \frac{a-1}{2X_t}\,dt + dB_t$$

(with $\mathbf{B}$ now a one-dimensional Brownian motion). For which values of $a \in \mathbf{R}$ does this define a homogeneous diffusion on $]l, r[=]0, \infty[$?

Example. The *Cox-Ingersoll-Ross* process is the homogeneous diffusion $\mathbf{X}$ on $]l, r[=]0, \infty[$ solving

$$dX_t = (a + bX_t)\,dt + \sigma\sqrt{X_t}\,dB_t.$$

We must of course investigate for which values of $a, b \in \mathbf{R}$ and $\sigma > 0$ (i) and (ii) from the *MAIN THEOREM* are satisfied. But

$$\begin{aligned} S'(x) &= x^{-\frac{2a}{\sigma^2}} e^{-\frac{2b}{\sigma^2}x}, \\ k(x) &= \frac{2}{\sigma^2} x^{\frac{2a}{\sigma^2}-1} e^{\frac{2b}{\sigma^2}x}. \end{aligned}$$

An effort evaluating various integrals reveals that (i) and (ii) are satisfied iff

$$2a \geq \sigma^2.$$

With this satisfied, $S(0) = -\infty$ always and $S(\infty) = \infty$ iff either $b < 0$ or $b = 0$, $2a = \sigma^2$. Thus $\mathbf{X}$ is recurrent iff $b < 0$ or $b = 0$, $2a = \sigma^2$, and in all other cases $P^x(\lim X_t = \infty) = 1$.

Exercise. Please carry out the effort required for evaluating the integrals relevant for verifying (i) and (ii) for the Cox-Ingersoll-Ross process.

STATIONARITY

Let $\mathbf{X}$ be a one-dimensional homogeneous diffusion on $]l,r[$ with continuous drift b and continuous diffusion function $\sigma > 0$, satisfying (i) and (ii) from the *MAIN THEOREM*. Suppose that X_0 is an $\boldsymbol{F}_0$-measurable random variable, independent of $\mathbf{B}$. Then $\mathbf{X}$ is *stationary* if all X_t have the same distribution (that of X_0) and the common distribution μ is the *stationary initial distrbution for* $\mathbf{X}$. When does such a μ exist and how is it found?

If μ exists and X_0 has distribution μ, because $\mathbf{X}$ is a Markov process, it is *strictly stationary:* for any $t > 0$, the process $(X_{t+u})_{u \geq 0}$ has the same distribution as $\mathbf{X}$ itself.

Suppose μ exists and that X_0 has distribution μ. For $f :]l,r[\to \mathbf{R}$ bounded with $Af = bf' + \frac{1}{2}\sigma^2 f''$ bounded,

$$f(X_t) = f(X_0) + \int_0^t Af\left(X_s\right)\, ds + M_t$$

with $\mathbf{M}$ a true martingale, $E^\mu M_t = 0$ for all t. But $E^\mu f(X_t) = E^\mu f(X_0)$ so

$$\begin{aligned} 0 &= E^\mu \int_0^t Af\left(X_s\right)\, ds \\ &= \int_0^t E^\mu Af\left(X_s\right)\, ds \\ &= t\mu\left(Af\right) \end{aligned}$$

writing

$$\mu\left(g\right) = \int_l^r g(x)\,\mu\left(dx\right).$$

Thus $\mu\left(Af\right) = 0$ for all $f \in C^2$ bounded with Af bounded. Guessing that μ has a nice density ϕ, take $f \in C^2$ such that $f' \equiv 0$ on $]l,a]$ and $[b,r[$ for some $a < b \in]l,r[$. Then

$$\begin{aligned} 0 &= \int_l^r \left(b(x)f'(x) + \frac{1}{2}\sigma^2(x)f''(x)\right)\phi(x)\,dx \\ &= \int_l^r f'(x)\left(b(x)\phi(x) - \frac{1}{2}\left(\sigma^2(x)\phi(x)\right)'\right)\,dx \end{aligned}$$

by partial integration. Let f' vary to obtain

$$b(x)\phi(x) = \frac{1}{2}\left(\sigma^2(x)\phi(x)\right)'$$

or

$$\frac{b(x)}{\sigma^2(x)}\left(\sigma^2(x)\phi(x)\right) = \frac{1}{2}\left(\sigma^2(x)\phi(x)\right)'.$$

Solve for

$$\begin{aligned} \sigma^2(x)\phi(x) &= \exp\left(\int_{x_1}^x \frac{2b(y)}{\sigma^2(y)}\,dy\right) \\ &= \frac{1}{S'(x)}. \end{aligned}$$

It follows that ϕ is proportional to the density k for the speed measure, in particular it must hold that $\int_l^r k(x)\,dx < \infty$. Note that then, since (i) and (ii) from the main theorem are satisfied, necessarily $\mathbf{X}$ is recurrent, $S(r) = \infty$, $S(l) = -\infty$. We have argued part of the following result:

THEOREM. The diffusion $\mathbf{X}$ on $]l, r[$ with drift b and diffusion function σ, has a stationary initial distribution μ iff the speed measure m satisfies

$$K = m\left(]l, r[\right) < \infty.$$

In that case

$$\mu = \frac{1}{K} m.$$

In order for μ to exist it is necessary (but not sufficient) that $\mathbf{X}$ be recurrent.

Example. $\mathbf{B}$ itself is *null-recurrent:* recurrent but no stationary initial distribution.

Example. The Ornstein-Uhlenbeck process on $]l, r[= \mathbf{R}$,

$$dX_t = -\rho X_t \, dt + dB_t$$

with $\rho \neq 0$ we know $\mathbf{X}$ to be recurrent iff $\rho > 0$. In that case there is a stationary initial distribution, the normal distribution with mean 0 and variance $\frac{1}{2\rho}$.

Example. The Cox-Ingersoll-Ross process on $]l, r[=]0, \infty[$,

$$dX_t = (a + bX_t) \, dt + \sigma \sqrt{X_t} \, dB_t$$

is well defined iff $2a \geq \sigma^2$ and recurrent iff in addition $b < 0$ or $b = 0$, $2a = \sigma^2$. If $b = 0$, $2a = \sigma^2$ the diffusion is null-recurrent. If $b < 0$, $2a \geq \sigma^2$ there is a stationary initial distribution, which is a Γ-distribution.

CHANGES OF MEASURE

Let $\mathbf{X}$ be a one-dimensional Brownian motion with drift $\xi \in \mathbf{R}$ and variance $\sigma^2 > 0$ and initial state $X_0 \equiv x_0 \in \mathbf{R}$, ie

$$X_t = x_0 + \xi t + \sigma B_t$$

with $\mathbf{B}$ a standard one-dimensional Brownian motion. Let $t_0 > 0$ and suppose that $\mathbf{X}$ is observed completely on the time interval $[0, t_0]$ — unrealistic in practice, but of great (!) theoretical interest. How should one estimate the parameters ξ and σ^2 from this complete observation of $\mathbf{X}$?

Fact. For $\mathbf{B}$ itself it holds that (with probability 1) simultaneously for all $t \geq 0$

$$\lim_{n \to \infty} \sum_{i=1}^{[2^n t]} \left(B_{\frac{i}{2n}} - B_{\frac{i-1}{2^n}} \right)^2 = t, \tag{3}$$

a result that carries over to $\mathbf{X}$ in the following form,

$$\lim_{n \to \infty} \sum_{i=1}^{[2^n t]} \left(X_{\frac{i}{2n}} - X_{\frac{i-1}{2^n}} \right)^2 = \sigma^2 t.$$

(3) expresses that the *quadratic variation* process for $\mathbf{B}$ is the deterministic function $t \mapsto t$.

This fact shows that one can estimate σ^2 precisely (!) from the complete observation of $\mathbf{X}$ on $[0, t_0]$, eg

$$\sigma^2 = \lim_{n \to \infty} \frac{1}{t_0} \sum_{i=1}^{[2^n t_0]} \left(X_{\frac{i}{2n}} - X_{\frac{i-1}{2^n}} \right)^2.$$

So it only remains to estimate ξ. Suppose for convenience that $\sigma^2 = 1$ and recall (?) that $\mathbf{Z}$ defined by

$$Z_t = \exp\left(\xi B_t - \frac{1}{2}\xi^2 t\right) \tag{4}$$

is a martingale with $EZ_t = 1$.

Let $Q_\xi^{x_0}$ denote the distribution of $\mathbf{X}$ (when $\sigma^2 = 1$), in particular $Q^{x_0} = Q_0^{x_0}$ is the distribution of the process $\mathbf{B} + x_0$. Recall that each $Q_\xi^{x_0}$ is a probability measure on the space $W = W^{(1)}$ of continuous functions $w : [0, \infty[\rightarrow \mathbf{R}$, equipped with the σ-algebra $\boldsymbol{H} = \boldsymbol{H}^{(1)}$. Defining $X_t^\circ : W \rightarrow \mathbf{R}$ by

$$X_t^\circ(w) = w(t),$$

$\boldsymbol{H}$ is the smallest σ-algebra of subsets of W that make all X_t° measurable. We then define the filtration $(\boldsymbol{H}_t)_{t \geq 0}$ with $\boldsymbol{H}_t$ the smallest σ-algebra that make all X_s° for $s \leq t$ measurable.

THEOREM. For all $t > 0$ and $H \in \boldsymbol{H}_t$,

$$Q_\xi^{x_0}(H) = \int_H \exp\left(\xi(X_t^\circ - x_0) - \frac{1}{2}\xi^2 t\right) \, dQ^{x_0}. \tag{5}$$

Note. The integrand on the right of course comes from (4). Because $EZ_t = 1$, for $H = W$ the integral $= 1$ and therefore, as a function of H defines a probability measure on W. To verify that this probability is $Q_\xi^{x_0}$ requires work!

The main message from (5) is that for the statistical problem of estimating ξ when $\mathbf{X}$ is completely observed on $[0, t_0]$ (and $\sigma^2 = 1$), as *likelihood function* we may use

$$\xi \mapsto \exp\left(\xi(X_{t_0} - x_0) - \frac{1}{2}\xi^2 t_0\right).$$

Maximising this yields the simple maximum-likelihood estimator

$$\widehat{\xi} = \frac{1}{t_0}(X_{t_0} - x_0).$$

We shall consider a more general statistical estimation problem. Suppose given a family of SDEs for homogeneous one-dimensional diffusions on a given interval $]l, r[$,

$$dX_t = b_\theta(X_t)\, dt + \sigma(X_t)\, dB_t, \quad X_0 \equiv x_0$$

with θ an unknown parameter (one- or many-dimensional). It is assumed of course that all b_θ and σ are continuous with $\sigma > 0$ and that for all θ, (i) and (ii) from the *MAIN THEOREM* are satisfied. Note that the diffusion function σ *does not depend on* θ: this is the analogue to assuming variance σ^2 of the Brownian motion known in the discussion above.

Note. The quadratic variation process for $\mathbf{X}$ for any θ is

$$\left(\int_0^t \sigma^2(X_s)\, ds\right)_{t \geq 0}$$

which does not depend on θ.

Let $Q_\theta^{x_0}$ denote the distribution of $\mathbf{X}$ when the true parameter value is θ — a probability measure on W which should now be the space of continuous functions $w : [0, \infty[\rightarrow]l, r[$ with the definitions of X_t°, $\boldsymbol{H}$ and $\boldsymbol{H}_t$ adjusted accordingly. Fix an arbitrary value θ_0 of θ and consider the problem of finding for each θ a process $\mathbf{Z}_\theta^\circ$ (adapted to $(\boldsymbol{H}_t)$) defined on W such that for all $t > 0$, $H \in \boldsymbol{H}_t$,

$$Q_\theta^{x_0}(H) = \int_H Z_{\theta,t}^\circ \, dQ_{\theta_0}^{x_0}. \tag{6}$$

With $\mathbf{X}$ completely observed on $[0, t]$ we would then use

$$\theta \mapsto Z_{\theta,t}^\circ$$

as likelihood function and find the maximum-likelihood estimator by maximisation. For this procedure, the choice of θ_0 does not matter! (Why?)

The problem of whether the $\mathbf{Z}_\theta^\circ$ exist and what they look like is certainly non-trivial. One simple observation is the following: in order that the integral on the right of (6) define a probability when H varies, we must obviously have

$$\int_W Z_{\theta,t}^\circ \, dQ_{\theta_0}^{x_0} = 1$$

and $Z_{\theta,t}^\circ \geq 0$. Furthermore, since for $0 \leq s < t$ and $H \in \boldsymbol{H}_s$ we also have $H \in \boldsymbol{H}_t$ so

$$Q_\theta^{x_0}(H) = \int_H Z_{\theta,s}^\circ \, dQ_{\theta_0}^{x_0} = \int_H Z_{\theta,t}^\circ \, dQ_{\theta_0}^{x_0},$$

ie for every θ, $\mathbf{Z}_\theta^\circ$ *is a* $Q_{\theta_0}^{x_0}$*-martingale*!

THEOREM. With respect to the reference value θ_0, $\mathbf{Z}_\theta^\circ$ is given by

$$Z_{\theta,t}^\circ = \exp\left(\int_0^t \frac{(b_\theta - b_{\theta_0})(X_s^\circ)}{\sigma^2(X_s^\circ)} \, dX_s^\circ - \frac{1}{2}\int_0^t \frac{(b_\theta^2 - b_{\theta_0}^2)(X_s^\circ)}{\sigma^2(X_s^\circ)} \, ds\right)$$

The theorem is an example of a *Girsanov type theorem.* The expression for $Z_{\theta,t}$ is an example of a *Cameron-Martin formula.*

Warning. It is absolutely essential for the validity of the theorem that for *all* θ, $\mathbf{X}$ is a diffusion on $]l, r[$ with $\sigma(x) > 0$. Don't ever try to use it without verifying this! To be on the safe side, it also helps that b and σ are continuous.

LÉVY PROCESSES

A (one-dimensional) Lévy process $\mathbf{X}$ defined on $(\Omega, \boldsymbol{F}, \boldsymbol{F}_t, P)$ is an adapted cádlág process (right-continuous with left limits) with independent and stationary increments, ie for every $s \geq 0$, $t > 0$, $X_{t+s} - X_s$ is independent of $\boldsymbol{F}_s$ with a distribution μ_t that depends on t only.

The μ_t form a *convolution semigroup,* $\mu_{s+t} = \mu_s * \mu_t$ which is weakly continuous, ie

$$\lim_{t \downarrow 0} \int f(x) \, \mu_t(dx) = f(0)$$

for all bounded and continuous f.

According to the classical *Lévy-Kinchine formula,* the characteristic function for μ_t has the form

$$\begin{aligned} & Ee^{iu(X_{t+s} - X_s)} = \int e^{iux} \, \mu_t(dx) \\ = \;& \exp t \left\{ iub - \tfrac{1}{2}u^2\sigma^2 + \int_{\mathbf{R}\backslash 0} \left(e^{iux} - 1 - iux 1_{[-1.1]}(x)\right) \nu(dx) \right\}. \end{aligned}$$

Here $b \in \mathbf{R}$, $\sigma^2 \geq 0$ and ν, the *Lévy measure,* is a positive measure on $\mathbf{R}\backslash 0$ such that

$$\int_{[-\varepsilon,\varepsilon]\backslash 0} x^2 \, \nu(dx) < \infty, \quad \nu(\mathbf{R}\backslash[-\varepsilon,\varepsilon]) < \infty$$

for all $\varepsilon > 0$.

If $\nu \equiv 0$ and $X_0 \equiv 0$, $\mathbf{X}$ is Brownian motion with drift b and variance σ^2 (degenerating into $X_t \equiv bt$ if $\sigma^2 = 0$).

If

$$\int_{[-1,1]\backslash 0} |x| \; \nu(dx) < \infty$$

one may rewrite the Lévy-Kinchine formula in the simpler form

$$\int e^{iux}\,\mu_t(dx) = \exp t\left\{iu\widetilde{b} - \tfrac{1}{2}u^2\sigma^2 + \int_{\mathbf{R}\backslash 0}\left(e^{iux}-1\right)\,\nu\,(dx)\right\} \tag{7}$$

where

$$\widetilde{b} = b - \int_{[-1,1]\backslash 0} x\,\nu(dx).$$

Taking $\widetilde{b} = \sigma^2 = 0$ in (7) and $\nu = \lambda\varepsilon_1$, where $\lambda > 0$, μ_t becomes the Poisson distribution with parameter λt and if $X_0 \equiv 0$, $\mathbf{X} = \mathbf{N}$ is the homogeneous *Poisson process* with intensity parameter λ, $N_0 \equiv 0$. $\mathbf{N}$ increases in jumps of size 1, is constant between jumps and if $T_n = \inf\{t : N_t = n\}$ is the time of the n'th jump ($T_0 \equiv 0$), the waiting times $V_n = T_n - T_{n-1}$ between successive jumps are iid and exponentially distributed, $P(V_n > v) = e^{-\lambda v}$ for $v \geq 0$.

If $\widetilde{b} = \sigma^2 = 0$ in (7) and $\nu = \lambda\pi$ where $\lambda > 0$ and π is a probability measure on $\mathbf{R}\backslash 0$, the resulting $\mathbf{X}$ with $X_0 \equiv 0$ is a *compound Poisson process* starting from 0 that may also be described as follows:

$$X_t = \sum_{n=1}^{N_t} U_n$$

where $\mathbf{N}$ with $N_0 \equiv 0$ is Poisson with intensity λ and the U_n are iid with distribution π and independent of $\mathbf{N}$.

In general the Lévy measure ν completely describes the *jump structure* for $\mathbf{X}$. Introduce the process $\Delta\mathbf{X}$ of jumps,

$$\Delta X_t = X_t - X_{t-}$$

and for $B \subseteq \mathbf{R}\backslash 0$ define the process $\mathbf{N}(B)$ by

$$N_t(B) = \sum_{0<s\leq t} 1_{(\Delta X_s \in B)},$$

the number of jumps on $[0,t]$ of a size $\in B$. Write $\overline{\mathbf{N}} = \mathbf{N}(\mathbf{R}\backslash 0)$ for the process counting the total number of jumps.

THEOREM. (i) $\mathbf{X}$ is continuous iff $\nu \equiv 0$.
(ii) If $\nu(\mathbf{R}\backslash 0) < \infty$, $\overline{N}_t < \infty$ for all t.
(iii) If $\nu(\mathbf{R}\backslash 0) = \infty$, $\overline{N}_{s+t} - \overline{N}_s = \infty$ for all $s \geq 0$, $t > 0$ and the random set $\{t : \Delta X_t \neq 0\}$ of jump times is a countably infinite and dense subset of $]0,\infty[$.
(iv) If $B \subseteq \mathbf{R}\backslash 0$ with $0 < \nu(B) < \infty$, then $\mathbf{N}(B)$ is a Poisson process with intensity $\nu(B)$. If $B_1, \ldots, B_n \subseteq \mathbf{R}\backslash 0$ are pairwise disjoint with all $\nu\left(B_j\right) \in \,]0,\infty[$, then the Poisson processes $\mathbf{N}(B_1), \ldots, \mathbf{N}(B_n)$ are independent.
(v) If $\int_{[-1,1]\backslash 0} |x| \; \nu(dx) < \infty$, then for all t

$$\sum_{0<s\leq t} |\Delta X_s| < \infty.$$

(vi) If $\int_{]0,1]} x\,\nu(dx) = \infty$, ($\int_{[-1,0[} x\,\nu(dx) = -\infty$) then for all $s \geq 0$, $t > 0$

$$\sum_{s<r\leq s+t} \Delta X_r 1_{(\Delta X_r > 0)} = \infty, \quad \left(\sum_{s<r\leq s+t} \Delta X_r 1_{(\Delta X_r < 0)} = -\infty\right).$$

Fact. Given a weakly continuous convolution semigroup (μ_t) of probability measures on $\mathbf{R}$, the following recipe can be used for finding the Lévy measure:

$$\nu(f) = \lim_{t\downarrow 0} \frac{1}{t}\mu_t(f) \tag{8}$$

for all bounded and continuous functions f vanishing in a neighborhood of 0. (Notation: $\zeta(f) = \int f(x)\,\zeta(dx)$.)

Example. For the *Γ-process,* μ_t is the Γ-distribution with scale parameter $\alpha > 0$ and shape parameter βt with $\beta > 0$, ie μ_t is the probability on $]0, \infty[$ with density

$$\frac{1}{\alpha^{\beta t}\Gamma(\beta t)} x^{\beta t - 1} e^{-\frac{x}{\alpha}} \quad (x > 0).$$

Using (8) one finds that ν is the positive measure on $]0, \infty[$ with density

$$\frac{\beta}{x} e^{-\frac{x}{\alpha}} \quad (x > 0)\,.$$

We see that the Γ-process has infinitely many jumps on any (non-degenerate) interval and that the sum of the jumps is finite on any finite interval. In fact,

$$X_t = X_0 + \sum_{0<s\le t} \Delta X_s,$$

ie $\mathbf{X} - X_0$ is the sum of its jumps. Also (7) holds with ν the measure just found and $\widetilde{b} = \sigma^2 = 0$.

Example. For the *Cauchy process,* μ_t is the Cauchy distribution with scale parameter t, ie μ_t has density

$$\frac{1}{\pi t \left(1 + \frac{x}{t}\right)^2}$$

on $\mathbf{R}$. The Lévy measure has density

$$\frac{1}{\pi x^2}$$

on $\mathbf{R}\backslash 0$. So the Cauchy process has infinitely many jumps on any interval and on any interval the sum of the positive jumps and the sum of the negative jumps are both infinite. The heavy tails of the Lévy measure entails that the process often has huge jumps. Altogether, the Cauchy process is pretty wild!

It is perfectly possible to define new processes from Lévy processes by solving stochastic differential equations, typically of the form

$$dY_t = b(Y_t)\,dt + \sigma(Y_{t-})\;dX_t \tag{9}$$

where $\mathbf{X}$ is a Lévy process with $X_0 \equiv 0$.

It is important to write $\sigma(Y_{t-})$ instead of $\sigma(Y_t)$: if $\mathbf{X}$ has jumps, so will $\mathbf{Y}$ (if $\sigma(y) \neq 0$), in fact

$$\Delta Y_t = \sigma(Y_{t-})\,\Delta X_t,$$

ie

$$Y_t = Y_{t-} + \sigma(Y_{t-})\,\Delta X_t \tag{10}$$

so the value of Y_t is determined in a simple manner from the value just before the jump and the jump of $\mathbf{X}$.

A particularly simple case is that of *diffusions with (few) jumps:* in (9) let $\mathbf{X}$ be the sum of a standard Brownian motion $\mathbf{B}$ and an independent compound Poisson process. In between jumps (9) is just an SDE driven by $\mathbf{B}$, which we know how to solve, and at the jump times one uses (10) to obtain the solution.

This however exhibits one weakness of using SDEs driven by Lévy processes to construct new processes: if $\sigma(y) \neq 0$ for all y, $\mathbf{Y}$ will jump precisely when $\mathbf{X}$ does, so the jumps for $\mathbf{Y}$ arrive according to a Poisson process, which may be too restrictive a structure. To achieve a higher grade of flexibility one may combine SDEs of the form (9) with change of measure techniques.