LÉVY PROCESSES - LECTURE 1

ADAM ANDERSSON

1. Basic notions of Probability theory

A *probability space* $(\Omega, \mathcal{F}, \mathbf{P})$ is a measure space for which $\mathbf{P}(\Omega) = 1$. By a *random variable* we mean a measurable function

$$X\colon \Omega \to \mathbf{R},$$

We define the *law*, or *distribution* of X as the measure on $\mathbf{R}$, given by $\mathcal{L}(X) = P \circ X^{-1}$, where $X^{-1}(A)$ denotes the inverse image of A under X, for $A \in \mathcal{B}_{\mathbf{R}}$. Often this law is absolutely continuous with respect to the Lebesgue measure and we then refer to the Radon-Nikodym derivative as a *probability density function*. By the *expected value* of a random variable X we mean

$$\mathbf{E}[X] = \int_\Omega X(\omega)\,\mathrm{d}\mathbf{P}(\omega).$$

Often we allow ourself to write $\mathbf{E}X$ when no confusion arises. The *variance* of X is defined as

$$\mathrm{Var}(X) = \mathbf{E}[(X - \mathbf{E}X)^2] = \mathbf{E}[X^2] - \mathbf{E}[X]^2$$

and the *covariance* of two random variables X and Y as

$$\mathrm{Cov}(X, Y) = \mathbf{E}[(X - \mathbf{E}X)(Y - \mathbf{E}Y)].$$

Lebesgue spaces $L^q(\Omega)$ are defined as the equivalence classes of random variables X that satisfies $\mathbf{E}[|X|^q] < \infty$. Here X and Y belong to the same equivalence class if $\mathbf{E}[|X - Y|^q] = 0$.

A set $A \in \mathcal{F}$ is called an *event.* An event A is said to happen *almost surely* if $P(A) = 1$. A collection $(A_n)_{n\in\mathbf{N}} \subset \mathcal{F}$ of events are called *independent* if for any distinct $A_{i_1}, \ldots, A_{i_n}$

$$\mathbf{P}(A_{i_1}, \ldots, A_{i_n}) = \prod_{i=1}^n \mathbf{P}(A_{i_j}).$$

A family of stochastic variables $(X_n)_{n\in\mathbf{N}}$ is said to be independent if all the events $X_n^{-1}(A_n)$ are independent for all possible choices of $(A_n)_{n\in\mathbf{N}} \subset \mathcal{B}_{\mathbf{R}}$. Independence of $X, Y \in L_1(\Omega)$ implies that

$$\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[X].$$

By a *stochastic process* we mean a family of random variables $(X(t))_{t\in T}$, where T is some set, often $T = [0, \infty)$. In the latter case the *trajectory* $t \mapsto X(t)$ models random evolution in time. Two stochastic processes are called *modifications* of each others if

$$\mathbf{P}(X(t) \neq Y(t)) = 0, \quad \text{for all } t \in T.$$

We follow the probabilistic convention to call the Fourier transform of the probability law its characteristic function. Let X be a random variable and ν its law. Its characteristic function is

$$\phi(u) = \mathbf{E}[e^{iuX}] = \int_{\mathbf{R}} e^{iux}\,\mathrm{d}\nu(x).$$

2. Brownian motion

Let $T = [0, \infty)$. A stochastic process $(B_t)_{t \in T}$ on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ is called a *Brownian motion* if

W1: B(0)=0, almost surely;
W2: B has independent and centered Gaussian increments with $\mathrm{Cov}(B(t), B(s)) = \min(s,t)$, for $s, t \in T$;
W3: B has continuous paths, almost surely.

There are several ways how to construct a Brownian motion. We will sketch one that takes as the starting point a sequence $(\xi_n)_{n \in \mathbf{N}}$ of independent standard normal random variables on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. It is a sketch since we evoke two deep theorems, both due to Kolmogorov, without proof. We choose this approach since we will get the Wiener integral for free out of it. This is the Itô integral with integrands not depending on $\omega \in \Omega$. Let $L_2(\Omega) = L_2(\Omega, \mathcal{F}, \mathbf{P}; \mathbf{R})$ be the space of real valued square integrable random variables. The sequence $(\xi_n)_{n \in \mathbf{N}}$ is an orthonormal set of $L_2(\Omega)$. Indeed, by the independence and the zero mean $\mathbf{E}\xi_i\xi_j = \mathbf{E}\xi_i\mathbf{E}\xi_j = 0$, for $i \leq j$ and since ξ_i has unit variance $\mathbf{E}|\xi_i|^2 = 1$, for all $i \in \mathbf{N}$. Define the mapping

$$I \colon L_2(T) \to L_2(\Omega), \quad f \mapsto \sum_{n \in \mathbf{N}} \xi_n \langle e_i, f \rangle_{L_2(T)},$$

where $(e_n)_{n \in \mathbf{N}} \subset L_2(T)$ is an ON-basis. It is clearly well defined and is also an isometry since, for $f \in L_2(T)$, by Parseval's identity

(2.1)
$$\mathbf{E}|I(f)|^2 = \mathbf{E}\Big| \sum_{n \in \mathbf{N}} \xi_n \langle e_n, f \rangle_{L_2(T)} \Big|^2 = \sum_{n \in \mathbf{N}} \mathbf{E}|\xi_n|^2 |\langle e_n, f \rangle_{L_2(T)}|^2 = \sum_{n \in \mathbf{N}} |\langle e_n, f \rangle_{L_2(T)}|^2 = \|f\|^2_{L_2(T)}.$$

More generally, for $f, g \in L_2(T)$, by an application of the dominated convergence theorem

(2.2)
$$\begin{aligned} \mathbf{E}[I(f)I(g)] &= \mathbf{E} \sum_{m,n \in \mathbf{N}} \xi_m \xi_n \langle e_m, f \rangle_{L_2(T)} \langle e_n, g \rangle_{L_2(T)} = \sum_{m,n \in \mathbf{N}} \mathbf{E}\xi_m \xi_n \langle e_m, f \rangle_{L_2(T)} \langle e_n, g \rangle_{L_2(T)} \\ &= \sum_{n \in \mathbf{N}} \langle e_n, f \rangle_{L_2(T)} \langle e_n, g \rangle_{L_2(T)} = \Big\langle \sum_{n \in \mathbf{N}} \langle e_n, f \rangle_{L_2(T)} e_n, g \Big\rangle_{L_2(T)} = \langle f, g \rangle_{L_2(T)} \end{aligned}$$

The random variable $I(f)$ is for every $f \in L_2(T)$, as an $L_2(\Omega)$-limit of normal random variables, normal and from (2.1),

$$I(f) \sim N(0, \|f\|_{L_2(T)}), \quad \forall f \in L_2(T). \tag{2.3}$$

Let $\mathbf{1}_{[0,t]}$ be the indicator function that takes the value 1 in $[0,t]$ and is 0 in (t, ∞). We claim that $B(t) = I(\mathbf{1}_{[0,t]})$ is a Brownian motion. Clearly **W1** is satisfied since $\mathbf{1}_{[0,0]} = 0$ implies $B(0) = I(\mathbf{1}_{[0,0]}) = 0$, for all $\omega \in \Omega$. We verify **W2**, using the isometry (2.2). For $s, t \in T$

$$\mathrm{Cov}(B(s), B(t)) = \mathbf{E}B(s)B(t) = \mathbf{E}I(\mathbf{1}_{[0,s]})I(\mathbf{1}_{[0,t]}) = \langle \mathbf{1}_{[0,s]}, \mathbf{1}_{[0,t]} \rangle = \int_0^{\min(s,t)} \mathrm{d}r = \min(s,t).$$

So far it has been easy for us. For the continuity **W3** we need the following theorem. Its proof will not be presented since our objective further on is to study processes with jumps.

Theorem 2.1 (Kolmogorov's continuity criterion)**.** *Let $(X(t))_{t \in T}$ be a real valued stochastic process such that there exists positive constants α, β and C such that*

$$\mathbf{E}(|X(t_2) - X(t_1)|^\alpha) \leq C|t_2 - t_1|^{1+\beta} \tag{2.4}$$

for all $t_1, t_2 \in T$. Then there exists another stochastic process $(\tilde{X}(t))_{t \in T}$ such $\tilde{X}(t) = X(t)$ almost surely, for all $t \in T$, and such that the paths $t \mapsto \tilde{X}(t)$ is, almost surely, Hölder continuous with exponent γ, for every $\gamma \in (0, \frac{\beta}{\alpha})$.

We prove **W3** for B, by verifying (2.4) with constants $\alpha = 2n$, $\beta = n - 1$ and $C_n > 0$. From the definition it follows that $f \mapsto I(f)$ is linear. Thus, $B(t) - B(s) = I(1_{(s,t]})$ and hence

$B(t) - B(s) \sim N(0, \|1_{(s,t]}\|_{L_2(T)}) = N(0, t-s)$ from (2.1). We calculate the $2n$th moments, for $n \in \mathbf{N}$ using this fact. First

$$\mathbf{E}|B(t) - B(s)|^{2n} = \int_{\mathbf{R}} x^{2n} e^{-\frac{x^2}{2(t-s)}} \frac{\mathrm{d}x}{\sqrt{2\pi(t-s)}}$$

and by the change of variable $x = y\sqrt{t-s}$ we get that

$$\mathbf{E}|B(t) - B(s)|^{2n} = \int_{\mathbf{R}} y^{2n} e^{-\frac{y^2}{2}} \frac{\mathrm{d}y}{\sqrt{2\pi}} (t-s)^n = (2n-1)!\,(t-s)^n.$$

Thus there exist a process $(\tilde{B}(t))_{t\in T}$ that is a modification of $(B(t))_{t\in T}$, in the sense of Theorem 2.1, that satisfies **W1-W3**, and has almost surely Hölder continuous paths with exponent γ, for every $\gamma \in (0, \frac{1}{2})$. We denote the continuous modification $\tilde{B}$ by B from now on and let it be our Brownian motion. We conclude with a theorem

Theorem 2.2. *The paths $t \mapsto B(t)$ of a Brownian motion B are almost surely Hölder continuous, for any exponent $\gamma \in (0, \frac{1}{2})$, i.e.,*

$$\mathbf{P}\Big(\sup_{s\neq t} \frac{|B(t) - B(s)|}{|t-s|^\gamma} \leq \delta\Big) = 1$$

for some $\delta > 0$ large enough.

Theorem 2.3. *There exist a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ that admits the existence of Brownian motion.*

Proof. All we need is the existence of a probability space admitting a countable sequence of independent standard normal random variables. Let $\Omega = \mathbf{R}^\infty = \{(\omega_1, \omega_2, \omega_3, \dots)\colon \omega_n \in \mathbf{R}\}$. By $\Lambda \subset \mathcal{P}(\mathbf{N})$ we denote the collection of sets $A = (a_1, \dots, a_n) \subset \mathbf{N}$, $n \in \mathbf{N}$, that has cardinality $|A| < \infty$. For $A \in \Lambda$ we denote by π_A the projection $\mathbf{R}^\infty \to \mathbf{R}^{|A|}$ given by

$$\omega \mapsto (\omega_{a_1}, \dots, \omega_{a_{|A|}}),$$

where $a_1, \dots, a_{|A|} \in A$ is taken in increasing order. Let $\mathcal{C}$ be the algebra (closed under finite unions and intersections) of all subsets of $\mathbf{R}^\infty$, that are obtained as the inverse image $\pi_A^{-1}(B)$ of $B \in \mathcal{B}(\mathbf{R}^{|A|})$, i.e.,

$$\mathcal{C} = \{\pi_A^{-1}(B)\colon A \in \Lambda, B \in \mathcal{B}(\mathbf{R}^{|A|})\}.$$

All random variables of the form $F = f(\omega_{i_1}, \dots, \omega_{i_n})$ are measurable with respect to $\mathcal{C}$. It makes sense to choose the σ-algebra $\mathcal{F} = \sigma(\mathcal{C})$, generated by $\mathcal{C}$ for Ω. With respect to $\mathcal{F}$ all random variables $F = f(\omega_1, \omega_2, \dots)$, depending on the whole range of ω_n, $n \in \mathbf{N}$ are measurable.

We define a family of finite dimensional distributions, indexed by Λ, by $\mathbf{P}_A = \gamma^{|A|}$, where

$$\mathrm{d}\gamma^{|A|}(x) = (2\pi)^{-\frac{|A|}{2}} e^{-\frac{|x|^2}{2}}\,\mathrm{d}x.$$

Kolmogorov's extension theorem, see Theorem 10.18 in [2] or Theorem 2.2, p.50 in [4], guarantees the existence of of a probability measure $\mathbf{P}$ that agrees with the finite dimensional distributions, in the sense that $P \circ \pi_A^{-1} = P_A$. It will be stated and proved during the next lecture. $\square$

Those already familiar with stochastic analysis are perhaps amused to see a very simple proof of the Cameron-Martin theorem. Those how have never seen this theorem before will not suffer by proceed to the next section. A counterpart of this theorem, for Lévy processes is given in Chapter 5 of [1] but will not be treated in the course.

Theorem 2.4 (The Cameron-Martin Theorem)**.** *Let $(B(t))_{t\in T}$ be a Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Then, for $f \in L_2(T)$ the process*

$$\tilde{B}(t) = B(t) + \int_0^t f(s)\,ds$$

is a Brownian motion under the probability measure $\mathbf{Q}$, having Radon-Nykodym derivative

$$\frac{d\mathbf{Q}}{d\mathbf{P}} = \exp\Big(I(f) - \frac{1}{2}\|f\|^2\Big)$$

Proof. Using that $B(t) = I(\mathbf{1}_{[0,t]})$ and a Fourier expansion of f yields

$$\begin{aligned}\tilde{B}(t) &= I(\mathbf{1}_{[0,t]}) + \langle f, \mathbf{1}_{[0,t]}\rangle \\ &= \sum_{n\in\mathbf{N}} \xi_n \langle e_n, \mathbf{1}_{[0,t]}\rangle + \Big\langle \sum_{n\in\mathbf{N}} \langle e_n, f\rangle e_n, \mathbf{1}_{[0,t]}\Big\rangle \\ &= \sum_{n\in\mathbf{N}} [\xi_n + \langle e_n, f\rangle]\langle e_n, \mathbf{1}_{[0,t]}\rangle.\end{aligned}$$

If we can change the measure so that $\xi_n + \langle e_n, f\rangle \sim N(0,1)$, for each $n \in \mathbf{N}$, then $\tilde{B}$ is a Brownian motion with respect to the new measure. The probability density function of each $\xi_n + \langle e_n, f\rangle$ is

$$\frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{(x-\langle e_n, f\rangle)^2}{2}\Big) = \frac{1}{\sqrt{2\pi}} \exp\Big(\frac{2x\langle e_n, f\rangle - |\langle e_n, f\rangle|^2}{2}\Big) \exp\Big(-\frac{x^2}{2}\Big).$$

Thus, under the measure

$$\exp\Big(\xi_n\langle e_n, f\rangle - \frac{1}{2}|\langle e_n, f\rangle|^2\Big)\,\mathrm{d}\mathbf{P}$$

we have $\xi_n + \langle e_n, f\rangle \sim N(0,1)$ and with

$$\frac{\mathrm{d}\mathbf{Q}}{\mathrm{d}\mathbf{P}} = \prod_{n\in\mathbf{N}} \exp\Big(\xi_n\langle e_n, f\rangle - \frac{1}{2}|\langle e_n, f\rangle|^2\Big) = \exp\Big(I(f) - \frac{1}{2}\|f\|^2\Big)$$

$\xi_n + \langle e_n, f\rangle \sim N(0,1)$ under $\mathbf{Q}$, for every $n \in \mathbf{N}$. □

3. The Wiener integral

In this section a first example of a stochastic integral is presented, the Wiener integral. For that purpose consider a simple function

$$f = \sum_{n=1}^{N-1} a_n \mathbf{1}_{[t_n,t_{n+1})},$$

where $0 \le t_1 < t_2 < \ldots t_N < \infty$ and $a_1, \ldots, a_N \in \mathbf{R}$. Then, by the linearity of I

$$I(f) = \sum_{n=1}^{N-1} a_n I(\mathbf{1}_{[t_n,t_{n+1})}) = \sum_{n=1}^{N-1} a_n (B(t_{n+1}) - B(t_n)).$$

It therefor makes sense to write it as a Riemann-Stiltjes integral with f as integrand and B as the integrator, i.e.,

$$I(f) = \int_0^\infty f(t)\,\mathrm{d}B(t).$$

This integral is known as the Wiener integral. Notice that it is perfectly well defined for all $f \in L_2(T)$ since I is. No limiting argument is needed. One can prove that Brownian motion has, almost surely, infinite variation and is therefore not a suitable integrator in the usual sense. We state some important properties of the Wiener integral.

Theorem 3.1. *For all $f, g \in L_2(T)$ and $\alpha, \beta \in \mathbf{R}$ the following properties for the Wiener intgral holds:*

Linearity: $\int_0^\infty (\alpha f(t) + \beta g(t))\,dB(t) = \alpha \int_0^\infty f(t)\,dB(t) + \beta \int_0^\infty g(t)\,dB(t)$
Wiener's isometry: $\mathbf{E}\Big[\int_0^\infty f(t)\,dB(t) \int_0^\infty g(t)\,dB(t)\Big] = \langle f, g\rangle_{L_2(T)}$
Normality: $\int_0^\infty f(t)\,dB(t) \sim N(0, \|f\|^2_{L_2(T)})$*, for all* $f \in L_2(T)$

Proof. The linearity is obvious. The Wiener Isometry is nothing but (2.2) and the normality (2.3). □

Remark 3.2. We have developed a theory of stochastic integration with respect to deterministic integrands. When the integrand is a stochastic process, with the suitable integrability conditions, the theory is much more involved. The stochastic integral is then called the Itô integral and we refer the reader to books on this subject, see for instance [4]. Since the above construction is a special case of the Itô theory this simple construction is seldom presented, but still very instructive.

4. Poisson processes

A random varable X with probability density function $f(t) = \lambda e^{-\lambda t}$, for $\lambda > 0$, is called exponential with intensity λ, written

$$X \sim \exp(\lambda).$$

It has mean $\mathbf{E}X = \lambda^{-1}$ and variance $\mathrm{Var}(X) = \lambda^{-2}$. The exponential distribution has the following nice property:

$$\mathbf{P}(X \in [t, t+h] | X \geq t) = \frac{\lambda \int_t^{t+h} e^{-\lambda s}\,\mathrm{d}s}{\lambda \int_t^{\infty} e^{-\lambda s}\,\mathrm{d}s} = \frac{e^{-\lambda t} - e^{-\lambda(t-h)}}{e^{-\lambda t}} = 1 - e^{-\lambda h} = \lambda h + o(h), \tag{4.1}$$

for all $t \geq 0$. This is often referred to as the lack of memory property. If X is the time until a light bulb breaks, and we want to predict its future life time after a time $t > 0$, then it only matter if it is working at the time t; its further history is irrelevant.

Let $(X_n)_{n\in\mathbf{N}}$ be a sequence of independent and exponentially distributed random variables with common intensity λ and let $T_n = \sum_{i=1}^n X_i$. Let $(N(t))_{t\geq 0}$ be the stochastic process that satisfies $N(0) = 0$, stays constant a time T_1 and jump to $N(T_1) = 1$, stays constant a time X_2 and make a jump $N(T_2) = N(X_1 + X_2) = 2$ and so on. We make the formal definition:

$$N(t) = \max\{n \colon T_n \leq t\}.$$

The process $(N(t))_{t\geq 0}$ is called a *Poisson process* with intensity $\lambda > 0$.

A discrete random variable Y is called Poisson distributed, with parameter $\lambda > 0$, if it satisfies

$$\mathbf{P}(Y = n) = \frac{\lambda^n}{n!} e^{-\lambda}, \quad n = 0, 1, 2, \ldots$$

One can prove that $N(t)$ is Poisson distributed with parameter λt. For more about Poisson processes see [3].

5. Lévy processes

A process $(X(t))_{t\geq 0}$ is said to have stationary increments if for every $0 < s < t$, the increment $X(t) - X(s)$ has the same law as $X(t-s) - X(0)$. A stochastic process $(L(t))_{t\geq 0}$ is called a Lévy process if

L1: L(0)=0, almost surely;
L2: L has independent and stationary increments;
L3: L is stochastically continuous, in the sense that, for all $a > 0$ and for all $s \geq 0$

$$\lim_{t\to s} \mathbf{P}(|X(t) - X(s)| > a) = 0.$$

The notion of stochastic continuity is a rather weak one. It includes processes with jumps but tells that the probability of a jump at any given time $t \geq 0$ is zero.

Example 5.1. Brownian motion is a Lévy process. Indeed, **W1** and **L1** are the same, **W2** is a special case of **L2** and obviously the continuity **W3** is stronger than the notion of stochastic continuity of **L3**.

Example 5.2. The linear deterministic process $L(t) = ct$, for $c > 0$ is a Lévy process.

Example 5.3. The Poisson process $(N(t))_{t\geq 0}$, with intensity $\lambda > 0$, is a Lévy process. First, **L1** is a part of the definition. For $0 < s < t$, the process $(N(t) - N(s))_{t\geq s}$ is a new Poisson process with the same λ, except for the starting time. This holds due to the lack of memory property of exponential random variables. More precisely the time to the first jump of $N(t) - N(s)$ is

$\exp(\lambda)$ distributed, even if time has elapsed since the jump just prior to s. Next, $N(t) - N(s)$ has the same distribution as that of $N(t-s)$, for all $s < t$, since that too, of cause, is a Poisson process with the same intensity. Therefore **L2** is satisfied. For **L3** we notice, from (4.1), that the following holds:

$$\mathbf{P}(N(t+h) = n + m \,|N(t) = n) = \begin{cases} \lambda h + o(h), & \text{if } m = 1, \\ o(h), & \text{if } m > 1, \\ 1 - \lambda h + o(h), & \text{if } m = 0. \end{cases}$$

For every $n \in \mathbf{N}$, $m \geq 1$ this probability tends to 0 as we let $h \to 0$. This is the stochastic continuity.

References

[1] D. Applebaum *Lévy Processes and Stochastic Calculus*, Cambridge **2009**.
[2] G.B. Folland *Real Analysis: Modern Techniques and Their Applications,* Wiley **1999**
[3] G. Grimmet and D. Stirzaker *Probability and random processes,* Oxford **2001**
[4] I. Karatzas and S.E. Shreve *Brownian Motion and Stochastic Calculus,* Springer **1991**