2

Brownian motion

This second lecture is devoted to introduce two main subjects in probability theory: Gaussian measures and Brownian motion. In the first section, we shall introduce the class of Gaussian probability measures on $E = \mathbb{R}^n$, while the second section will be devoted to study Gaussian processes. Both these classes of objects are fundamental in modern probability theory and shall return frequently in the following.

There are several technical points from abstract measure theory involved in the definition of Browian motion. In this lecture, we shall see Brownian motion as an special case of Gaussian process. This characterization will provide useful in the proof of existence of this process, which is the object of third section. We shall appeal to some fundamental theorems due to Kolmogorov in order to prove the result. First, Kolmogorov's existence theorem 2.13 proves the existence of a stochastic process with the same finite dimensional distributions as Brownian motion; this does not suffice, since our definition requires continuity of sample paths: but finally we can appeal to Kolmogorov's theorem of continuity, Theorem 2.16, and conclude the proof of existence.

2.1 Gaussian measures

We start with the simple one dimensional case $E = \mathbb{R}$; we shall say that a Borel probability measure μ on $\mathbb{R}$ is a *Gaussian law* (or a *Gaussian distribution*) $\mathcal{N}(a, \sigma^2)$, for $a \in \mathbb{R}$, $\sigma^2 > 0$, if its density with respect to the Lebesgue measure exists and is given by the function

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-a)^2\right), \quad x \in \mathbb{R}.$$

If $\sigma^2 = 0$ we set

$$\mathcal{N}(a, 0) = \delta_a,$$

where δ_a is the Dirac measure (concentrated) in a. If $a = 0$, then $\mathcal{N}(0, \sigma^2)$ is called a centered Gaussian measure.

Let $X : \Omega \to \mathbb{R}$ be a random variable; we say that X is a *Gaussian random variable* if its law μ is a Gaussian measure; in particular, if $\mu = \mathcal{N}(a, \sigma^2)$ we write $X \sim \mathcal{N}(a, \sigma^2)$. In this case, X has mean $\mathbb{E}[X] = a$ and variance $\mathrm{Var}(X) = \sigma^2$. The characteristic function of X (i.e., the Fourier transform of μ) is given by

$$\hat{\mu}(\theta) = \mathbb{E}[\mathrm{e}^{\mathrm{i}X\theta}] = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}\theta x} \mu(\mathrm{d}x) = \mathrm{e}^{\mathrm{i}\theta a} \mathrm{e}^{-\frac{1}{2}\sigma^2\theta^2}.$$

Problem 2.1. Given a sequence $\{\sigma_n,\ n \in \mathbb{N}\}$ of real numbers, converging to 0, let $\{X_n\}$ be a sequence of centered Gaussian random variables, each with variance σ_n^2. Show that the sequence $\{X_n\}$ converges in law to Dirac's measure in 0.

Recall that for given Borel probability measures on $\mathbb{R}$ μ and ν, the convolution measure $\mu * \nu$ is defined by

$$(\mu * \nu)(B) = \int_{\mathbb{R}} \mu(B - x)\nu(dx), \qquad B \in \mathcal{B}(\mathbb{R}).$$

It is known that Fourier transform maps convolution of measures into pointwise product of characteristic functions; using this fact, and the above expression of characteristic function for a Gaussian law, we obtain easily the following result.

Proposition 2.1. *Given two Gaussian laws $\mu = \mathcal{N}(a, \sigma^2)$ and $\nu = \mathcal{N}(b, \tau^2)$, it holds*

$$\widehat{(\mu * \nu)}(\theta) = \hat{\mu}(\theta)\hat{\nu}(\theta) = \mathrm{e}^{\mathrm{i}\theta(a+b)} \exp\left(-\frac{1}{2}(\sigma^2 + \tau^2)\theta^2\right).$$

Hence, $\mu * \nu$ is the Gaussian law $\mathcal{N}(a + b, \sigma^2 + \tau^2)$. In particular, if X and Y are real independent Gaussian random variables, also $X + Y$ is a Gaussian random variable.

We consider next the multidimensional case. We are given n random variables $X_1, \ldots, X_n$, independent, with Gaussian law $\mathcal{N}(0, 1)$; setting $X = (X_1, \ldots, X_n)$, the law of the random vector X has density

$$f(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x_1^2/2} \ldots \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x_n^2/2} = \frac{1}{(2\pi)^{n/2}} \mathrm{e}^{-\frac{1}{2}|x|^2} \tag{2.1}$$

and characteristic function

$$\mathrm{e}^{-\frac{1}{2}\theta_1^2} \cdots \mathrm{e}^{-\frac{1}{2}\theta_n^2} = \mathrm{e}^{-\frac{1}{2}|\theta|^2}.$$

If Q is a matrix in $L^+(\mathbb{R}^n)$ (the space of symmetric non-negative defined $n \times n$ matrices), it can be expressed as $Q = AA^*$, where A is a $n \times n$ matrix and A^* denotes the adjoint matrix of A; moreover A is uniquely determined in the class of symmetric matrices: in this case, A is also denoted as $Q^{1/2}$. For given $a \in \mathbb{R}^n$ and $Q \in L^+(\mathbb{R}^n)$, then $Q^{1/2}X + a$ has characteristic function

$$\phi(\theta) = \mathrm{e}^{\mathrm{i}\langle\theta, a\rangle} \mathrm{e}^{-\frac{1}{2}\langle Q^{1/2}\theta, Q^{1/2}\theta\rangle} = \mathrm{e}^{\mathrm{i}\langle\theta, a\rangle} \mathrm{e}^{-\frac{1}{2}\langle Q\theta, \theta\rangle}, \quad \theta \in \mathbb{R}^n. \tag{2.2}$$

A Borel probability measure μ on $\mathbb{R}^n$ is a *(multidimensional) Gaussian law* $\mathcal{N}(a, Q)$ if its characteristic function is given by (2.2).

Remark that if the matrix Q is invertible then also $Q^{1/2}$ is invertible, and so there exists a density g for $Z = Q^{1/2}X + a$; this density is given from (2.1) by a change-of-variables formula

$$g(y) = \frac{1}{(2\pi)^{n/2}(\mathrm{Det}(Q))^{1/2}} \mathrm{e}^{-\frac{1}{2}\langle Q^{-1}(y-a), (y-a)\rangle}.$$

Conversely, if the covariance matrix Q is singular, $Q^{1/2}$ shall be singular as well, and the law of $Z = Q^{1/2}X + a$ is supported by some proper affine subspace of $\mathbb{R}^n$, so it cannot have a density. As a special case, we mention the law $\mathcal{N}(a, 0)$, Dirac's delta measure concentrated in a, having characteristic function $\theta \to \mathrm{e}^{\mathrm{i}\langle\theta,a\rangle}$.

The following result can be easily proved by using the Fourier transform.

Proposition 2.2. *Let X be a Gaussian random variable in $\mathbb{R}^n$ with distribution $\mathcal{N}(a, Q)$; consider the random variable $Y = AX + b$, where $A \in L(\mathbb{R}^n, \mathbb{R}^k)$ and $b \in \mathbb{R}^k$. Then Y has Gaussian distribution $\mathcal{N}(Aa + b, AQA^*)$.*

Applying the previous result, we get that, for each Gaussian random variable Z with values in $\mathbb{R}^n$, $Z \sim \mathcal{N}(a, Q)$, there exists a random variable X such that $Z = a + Q^{1/2}X$ and $X \sim \mathcal{N}(0, I)$ (I denotes the identity matrix in $\mathbb{R}^n$). Computing the derivatives in 0 of the characteristic function, we get that a is the mean of Z and Q is the covariance matrix.

In general, two random variables $X_1, X_2 : \Omega \to \mathbb{R}$ are said *uncorrelated* if their covariance matrix is diagonal. Of course, if they are independent then they are also uncorrelated, but in general the converse does not hold. However, for Gaussian families the following holds:

$$\begin{aligned}&\text{if } (X_1, X_2) \text{ is a Gaussian family, then}\\&X_1 \text{ and } X_2 \text{ Gaussian are independent} \iff \text{they are uncorrelated.}\end{aligned} \tag{2.3}$$

In fact, if X_1 and X_2 are uncorrelated, then the characteristic function of the vector $X = (X_1, X_2)$ has the form $\phi_X(h) = \exp(\mathrm{i}\langle h, m\rangle - \frac{1}{2}\langle Qh, h\rangle)$ where Q is the diagonal covariance matrix. Hence ϕ_X is equal to the product $\phi_{X_1} \cdot \phi_{X_2}$, so that X_1 and X_2 are independent, see Problem 1.5.

More generally, if X_1 and X_2 are respectively $\mathbb{R}^d$-valued and $\mathbb{R}^n$-valued random variables with joint Gaussian distribution such that

$$q_{ij} = \mathbb{E}[(X_1^i - \mathbb{E}[X_1^i])(X_2^j - \mathbb{E}[X_2^j])] = 0, \qquad i = 1, \dots, d \;\; j = 1, \dots, n,$$

then they are independent.

A similar result holds for n random variables as well; for instance, if $(X_1, \dots, X_n)$ is a Gaussian family of real random variables, then they are independent if and only if their covariance matrix is diagonal.

2.2 Stochastic processes

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the reference probability space; in this section we mainly consider stochastic processes, taking values in the Euclidean space $E = \mathbb{R}^d$, $d \geq 1$, endowed with the Borel σ-field $\mathcal{E} = \mathcal{B}(E)$.

Recall that a stochastic process $X = \{X_t,\ t \in T\}$ with values in E is a collection of random variables $X_t : \Omega \to E$, $t \in T$. It can be equivalently defined as a measurable mapping $X : (\Omega, \mathcal{F}) \to (E^T, \mathcal{E}^T)$, where $\mathcal{E}^T$ is the minimal σ-field which makes measurable all the projections $\pi_t : E^T \to E$, $\pi_t(f) = f(t)$, $f \in E^T$. We can also see X as a collection of random variables taking values in the space of paths E^T, since $\{X_t(\omega),\ t \in T\}$ is just an ordinary function of time (for every ω).

A stochastic process $\{X_t,\ t \in T\}$ is called *continuous* or *almost surely continuous* if its trajectories (or paths) $t \mapsto X_t(\omega)$, $T \to E$, are continuous (respectively, if almost every trajectory is continuous).

Similar definitions hold for *right continuous*, *almost surely right continuous* stochastic processes, and so on.

Let $\{\mathcal{F}_t, t \in T\}$ be a *filtration*, i.e., a family of sub-σ-fields of $\mathcal{F}$ increasing in time

$$\text{if } s < t \text{ then } \mathcal{F}_s \subset \mathcal{F}_t.$$

Intuitively, "$\mathcal{F}_t$ contains all the information which are available up to time t", i.e., all the events whose occourence can be established up to time t.

We add to the filtration $\{\mathcal{F}_t,\ t \in T\}$ the σ-algebra

$$\mathcal{F}_\infty = \bigvee_{t \in T} \mathcal{F}_t.$$

Recall that the right hand side stands for the minimal σ-field which contains all $\mathcal{F}_t$; this is not the same as $\underset{t\in T}{\cup} \mathcal{F}_t$. Define, for every $t \in T$

$$\mathcal{F}_{t+} = \bigwedge_{t<s\in T} \mathcal{F}_s = \bigcap_{t<s\in T} \mathcal{F}_s.$$

We have clearly, for each $t \in T$, $\mathcal{F}_t \subset \mathcal{F}_{t+}$.

Definition 2.3. *The filtration $\{\mathcal{F}_t,\ t \in T\}$ is* right-continuous *if for every $t \in T$, $\mathcal{F}_t = \mathcal{F}_{t+}$.*

Assume we are given a stochastic process $X = \{X_t,\ t \in T\}$ and a filtration $\{\mathcal{F}_t,\ t \in T\}$ on $(\Omega, \mathcal{F}, \mathbb{P})$. We say that the process X is *adapted* to the filtration $\{\mathcal{F}_t, t \in T\}$ if, for any $t \in T$ fixed, the random variable X_t is $\mathcal{F}_t$-measurable on Ω; equivalently, we say that X is adapted to $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$.

Notice that it is always possible to construct a filtration with respect to which the process is adapted, by setting $\mathcal{F}_t^X = \sigma(X_s, s \le t)$; $\mathcal{F}_t^X$ is called the *natural filtration* of X. A stochastic process X is adapted to a filtration $\{\mathcal{F}_t,\ t \in T\}$ if and only if one has $\mathcal{F}_t^X \subset \mathcal{F}_t$, $t \in T$.

Let $X = \{X_t,\ t \in T\}$ and $Y = \{Y_t,\ t \in T\}$ be two stochastic processes on $(\Omega, \mathcal{F})$. If the random variables X_t coincide almost surely with Y_t, i.e.,

$$\text{for each } t \in T: \qquad \mathbb{P}\{X_t \neq Y_t\} = 0,$$

X is called a *version* (or a *modification*) of the process Y.

In general, this does not imply that if X is adapted to a filtration $\mathcal{F}_t,\ t \in T$, then also Y is adapted to the same filtration. It may happen, in fact, that the set of zero measure $\{X_t \neq Y_t\}$ does not belong to $\mathcal{F}_t$, so that Y_t is not $\mathcal{F}_t$-measurable.

To avoid this and other technical difficulties, we shall always require that the filtration satisfies in addition the so called *standard assumptions* , i.e.,

(1) the filtration is right continuous, that is, for any $t \in T$: $\mathcal{F}_t = \mathcal{F}_{t+}$;
(2) the filtration is complete, that is, for any $t \in T$, $\mathcal{F}_t$ contains all $\mathbb{P}$-*null sets.*

Remark that each filtration $\{\mathcal{F}_t,\ t \in T\}$ can be completed in order to satisfies hypothesis (2). Indeed, it is enough to enlarge each $\mathcal{F}_t$ by considering the smallest σ-algebra $\mathcal{F}_t \vee \mathcal{N}$ which contains $\mathcal{F}_t$ and the family $\mathcal{N}$ of all $\mathbb{P}$-null sets.

Given a process X, the completion of the natural filtration $\{\mathcal{F}_t^X,\ t \in T\}$ will be called the *completed natural filtration* of X and still denoted by $\{\mathcal{F}_t^X,\ t \in T\}$.

The completed natural filtration of any Brownian motion satisfies the technical hypothesis (1); we will return on this later on.

Definition 2.4. *A* filtered probability space *or a* stochastic basis *is the quadruplet* $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$*, where the filtration* $\{\mathcal{F}_t\}$ *verifies the standard assumptions.*

When we say that a stochastic process $X = \{X_t,\ t \in T\}$ *is defined on a stochastic basis* $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$*, (implicitly) we also require that it is adapted to* $\{\mathcal{F}_t\}$*.*

Brownian motion

In the following, we introduce the main object of our investigations, namely Brownian motion.

Definition 2.5. *A d-dimensional* Brownian motion $B = \{B_t,\ t \geq 0\}$ *defined on a filtered probability space* $(\Omega, \mathcal{F}, \{\mathcal{F}_t,\ t \geq 0\}, \mathbb{P})$ *is a continuous stochastic process such that*

1. $B_0 = 0$ *a.s.;*
2. *for every* $0 \leq s < t$ *the random variable* $B_t - B_s$ *is independent from* $\mathcal{F}_s$*;*
3. *for every* $0 \leq s < t$ *the random variable* $B_t - B_s$ *has Gaussian law* $N(0, (t-s)I_d)$*, where* I_d *is the (d-dimensional) identity matrix.*

We shall again underline that, as opposite to Poisson process, we have just given a formal definition of Brownian motion without providing a construction of it, and we shall return later to the proof of its existence.

Remark 2.6. In the following, if the filtration $\{\mathcal{F}_t\}$ is not explicitly mentioned, we can assume that we are using the completed natural filtration $\{\mathcal{F}_t^B\}$ of B. Remark that in `Protter` [Pr04] it is shown that the completed natural filtration of any Brownian motion is right continuous and so it satisfies the standard assumptions. A Brownian motion B defined on $(\Omega, \mathcal{F}, \{\mathcal{F}_t^B,\ t \geq 0\}, \mathbb{P})$ is also called a *natural Brownian motion*.

Problem 2.2. (a) Shows that if in Definition 2.5 we consider the completed natural filtration $\{\mathcal{F}_t^B\}$, then condition 2. is equivalent to the following one:

2'. for any choice of times $0 < t_1 < \cdots < t_m$,

$$B_{t_1},\ B_{t_2} - B_{t_1}, \ldots,\ B_{t_m} - B_{t_{m-1}}$$

are independent random variables.

This condition says the Brownian motion is a process with independent increments (compare with (1) in Lemma 1.17 on the Poisson process).

(b) Consider a Brownian motion $B = \{B_t,\ t \geq 0\}$ defined on $(\Omega, \mathcal{F}, \{\mathcal{F}_t,\ t \geq 0\}, \mathbb{P})$; let $\mathcal{G}$ be a σ-field independent from $\mathcal{F}_\infty$. Setting $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{G}$, show that B is a $\{\mathcal{G}_t\}$-Browian motion.

Remark 2.7. To conclude this introduction, we compute the moments of a real Brownian motion B_t for fixed t. Since B_t has a Gaussian distribution $\mathcal{N}(0,t)$, for every $n \in \mathbb{N}$ there exists a constant $C_n = \frac{(2n)!}{2^n n!}$ such that

$$\mathbb{E}|Z|^{2n}] = \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} x^{2n} \mathrm{e}^{-x^2/2t}\, dx = t^n \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} y^{2n} \mathrm{e}^{-y^2/2}\, dy = C_n t^n.$$

In particular, for $n = 1$ and $n = 2$, we have

$$\mathbb{E}[|B_t|^2] = t, \qquad \mathbb{E}[|B_t|^4] = 3t^2.$$

2.2.1 Gaussian processes

Definition 2.8. *A stochastic process $X = \{X_t,\, t \in T\}$ is called a* Gaussian process *if for every choice of times $t_1, \dots, t_m \in T$ and scalars $\gamma_1, \dots, \gamma_m \in \mathbb{R}$, the random variable $\gamma_1 X_{t_1} + \dots + \gamma_m X_{t_m}$ has Gaussian distribution.*

The following proposition expresses a different characterization of Gaussian processes.

Proposition 2.9. *$X = \{X_t,\, t \in T\}$ is a Gaussian process if and only if for every choice of times $t_1, \dots, t_m \in T$ the random vector $(X_{t_1}, \dots, X_{t_m})$ has a Gaussian distribution in $\mathbb{R}^{d\cdot m}$.*

Proof. We will use Proposition 2.1. First, assume that X is a Gaussian process and fix times $t_1, \dots, t_m \in T$. We introduce $\xi = (X_{t_1}, \dots, X_{t_m})$ and prove that ξ has a Gaussian distribution. Take $d = 1$ for simplicity. For every vector $\gamma = (\gamma_1, \dots, \gamma_m) \in \mathbb{R}^m$, the scalar product $\langle \gamma, \xi \rangle$ is a real Gaussian random variable with mean a and variance σ^2, so that its characteristic function verifies

$$\mathbb{E}[\mathrm{e}^{\mathrm{i}\langle \gamma, \xi\rangle \theta}] = \mathrm{e}^{\mathrm{i}a\theta - \sigma^2\theta^2/2}, \quad \theta \in \mathbb{R}. \tag{2.4}$$

Note that in particular each X_{t_i} is a Gaussian random variable. We can compute a and σ^2 in terms of the mean μ and the covariance matrix Γ of the random vector ξ:

$$a = \mathbb{E}[\langle \gamma, \xi \rangle] = \mathbb{E}[\sum_{j=1}^m \gamma_j X_{t_j}] = \sum_{j=1}^m \gamma_j\, \mathbb{E}[X_{t_j}] = \sum_{j=1}^m \gamma_j \mu_{t_j} = \langle \gamma, \mu \rangle,$$

and similarly

$$\sigma^2 = \mathbb{E}[(\langle \gamma, \xi \rangle - \langle \gamma, \mu \rangle)^2] = \mathbb{E}[\langle \gamma, \xi - \mu \rangle^2] = \mathbb{E}[\sum_{i,j=1}^m \gamma_i \gamma_j (X_{t_i} - \mathbb{E}[X_{t_i}])(X_{t_j} - \mathbb{E}[X_{t_j}]) = \langle \Gamma\gamma, \gamma \rangle.$$

By (2.4) with $\theta = 1$, we find that the characteristic function of ξ is

$$\hat{\mu}(\gamma) = \mathbb{E}[\mathrm{e}^{\mathrm{i}\langle \gamma, \xi \rangle}] = \mathrm{e}^{\mathrm{i}\langle \gamma, \mu \rangle - \frac{1}{2}\langle \Gamma\gamma, \gamma \rangle},$$

and we recognize the characteristic function of a Gaussian random variable.

Conversely, if $\xi = (X_{t_1}, \dots, X_{t_m})$ is a random vector, then for every vector $\gamma \in \mathbb{R}^m$ the mapping $\xi \mapsto \langle \gamma, \xi \rangle$ is a linear transformation hence, by Proposition 2.2, maps Gaussian distributions into Gaussian distributions. It follows that X is a Gaussian process.

□

Given two stochastic processes $X = \{X_t,\, t \in T\}$ and $Y = \{Y_t,\, t \in T\}$ on $(\Omega, \mathcal{F}, \mathbb{P})$, we can express their independence in terms of the generated filtrations: the process $\{X_t,\ t \in T\}$ is independent from $\{Y_t,\ t \in T\}$ if $\sigma(X_t,\, t \in T)$ is independent from $\sigma(Y_t,\, t \in T)$.

Proposition 2.10. *Consider two families $X = \{X_t,\, t \in T\}$ and $Y = \{Y_t,\, t \in T\}$ such that $X \cup Y$ is Gaussian. Then X is independent from Y if and only if*

$$\mathbb{E}[X_t Y_s] = \mathbb{E}[X_t]\, \mathbb{E}[Y_s], \tag{2.5}$$

for every $t, s \in T$.

Proof. From (2.5) it follows that for any $t, s \in T$, the Gaussian random variables X_t and Y_s are independent, see (2.3). By a similar criterium, thanks to the arbitrariness of t and s, we have that any two random vectors $(X_{t_1}, \dots, X_{t_n})$ and $(Y_{s_1}, \dots, Y_{s_m})$ are independent.

Now the thesis follows thanks to Dynkin's theorem 1.3. One uses that $\sigma(X) = \sigma(X_t,\, t \in T)$ and $\sigma(Y) = \sigma(Y_t,\, t \in T)$ are generated by the π-systems of events $\{X_{t_1} \in A_1, \dots, X_{t_n} \in A_n\}$ and $\{Y_{s_1} \in B_1, \dots, Y_{s_m} \in B_m\}$ respectively, and that these events are independent due to the independence of the random vectors $(X_{t_1}, \dots, X_{t_n})$ and $(Y_{s_1}, \dots, Y_{s_m})$.

□

Problem 2.3. Give the details of the proof of Proposition 2.10.

Hint: first fix $A_1, \dots, A_n \in \mathcal{B}(\mathbb{R}^d)$ and set $B = \{X_{t_1} \in A_1, \dots, X_{t_n} \in A_n\}$. Define

$$\Lambda = \{A \in \sigma(Y) \text{ such that } A \text{ is independent from } B\}.$$

Show that Λ is a λ-system. Since Λ contains the events $\{Y_{s_1} \in B_1, \dots, Y_{s_m} \in B_m\}$ one deduces that $\Lambda = \sigma(Y)$. How can we finish the proof?

Using the above definition, it is possible to prove that a Brownian motion is a Gaussian process. Choose real numbers $\alpha_1, \dots, \alpha_m$ and times $0 < t_1 < \dots < t_m$: the goal is to prove that $\alpha_1 B_{t_1} + \dots + \alpha_m B_{t_m}$ is a random variable with Gaussian law. For $m = 1$ the claim obviously holds. By induction, we suppose that the claim holds for $m - 1$; thus, we can write

$$\alpha_1 B_{t_1} + \dots + \alpha_m B_{t_m} = [\alpha_1 B_{t_1} + \dots + (\alpha_{m-1} + \alpha_m) B_{t_{m-1}}] + \alpha_m (B_{t_m} - B_{t_{m-1}})$$

and the conclusion follows since the sum of two Gaussian, independent random variables again has a Gaussian distribution.

We may characterize a Brownian motion as a Gaussian process in the following way.

Proposition 2.11. *Given a real Brownian motion $B = \{B_t\}$, then*

a) $B_0 = 0$ *a.s.;*
b) *for any choice of times $0 \le t_1 < t_2 < \dots < t_m$ the m-dimensional random variable $(B_{t_1}, \dots, B_{t_m})$ has a centered Gaussian law;*
c) $\mathbb{E}(B_t B_s) = s \wedge t$.

Conversely, if a continuous process $\{B_t\}$ verifies the above conditions a), b) and c), then it is a natural Brownian motion (i.e., it is a Brownian motion with respect to its completed natural filtration).

Proof. If B is a Brownian motion, condition a) is obvious, condition b) follows by the characterization of Gaussian processes stated in Proposition 2.9 and condition c) holds since, for $s \le t$, we have

$$\mathbb{E}[B_t B_s] = \mathbb{E}[(B_t - B_s)B_s] + \mathbb{E}(B_s^2) = s = s \wedge t.$$

Conversely, if a Gaussian process B satisfies conditions a), b) and c), then property 1. of Definition 2.5 holds; also, the law of $B_t - B_s$ is Gaussian thanks to condition b), it is centered since B_t and B_s are, and its variance is equal to

$$\mathbb{E}[(B_t - B_s)^2] = \mathbb{E}[B_t^2] - 2\mathbb{E}[B_t B_s] + \mathbb{E}[B_s^2] = t - 2t \wedge s + s = t - s,$$

hence $B_t - B_s$ has Gaussian law $N(0, t-s)$. Finally, to show that $B_t - B_s$ is independent from $\mathcal{F}_s^B$, thanks to Proposition 2.10 it is enough to prove that, for each $\tau \le s$, $\mathbb{E}[(B_t - B_s)B_\tau] = 0$:

$$\mathbb{E}[(B_t - B_s)B_\tau] = \mathbb{E}[B_t B_\tau] - \mathbb{E}[B_s B_\tau] = t \wedge \tau - s \wedge \tau = \tau - \tau = 0.$$

□

2.3 Kolmogorov's existence theorem

Often, a stochastic process $X = \{X_t,\ t \in T\}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is described by means of its *finite dimensional distributions*. These are defined as the probability measures $\mu_{t_1,\dots,t_m}$ that X induces on $\mathcal{B}(E^m)$, for each $m \in \mathbb{N}$ and any choice of times $t_1 < t_2 < \cdots < t_m$ in T:

$$\mu_{t_1,\dots,t_m}(A_1 \times \cdots \times A_m) = \mathbb{P}(X_{t_1} \in A_1, \dots, X_{t_m} \in A_m), \qquad A_1, \dots, A_m \in \mathcal{E}.$$

We say that two stochastic processes with values in E (even defined on different probability spaces) are *equivalent* or that they are *equal in law* if they have the same finite dimensional distributions.

Remark 2.12. Let us give a heuristic justification of the term *equality in law*. A stochastic process $X = \{X_t,\ t \in T\}$ can be seen as a random variable with values in E^T. Hence, we may define the law μ_X of X on E^T, endowed with the product σ-algebra $\mathcal{E}^T$, i.e., $\mu_X(F) = \mathbb{P}(X \in F)$, $F \in \mathcal{E}^T$. Now, since μ_X is completely determined by its values on the sets of the form

$$\{f \in E^T \mid f(t_1) \in A_{t_1}, \dots, f(t_m) \in A_{t_m}\}, \qquad A_{t_i} \in \mathcal{E}, i = 1, \dots, m,$$

we have that the finite dimensional distributions of X characterize the law μ_X.

If $\{X_t,\ t \in T\}$ is a real Gaussian process, then its (finite dimensional) distributions are determined by the *mean function*

$$m(t) = \mathbb{E}[X_t]$$

and the *covariance function*

$$\rho(t, s) = \mathbb{E}[(X_t - m(t))(X_s - m(s))].$$

Indeed, the joint density of $X_{t_1}, \dots, X_{t_n}$ is just

$$\mu_{t_1,\dots,t_n}(A_1 \times \dots \times A_n) = \mathcal{N}(m(\boldsymbol{t}), \Sigma)(A_1 \times \dots \times A_n) \tag{2.6}$$

where Σ is the symmetric $n \times n$ matrix

$$\Sigma = \Sigma(t_1, \dots, t_n) = \begin{pmatrix} \rho(t_1,t_1) & \rho(t_1,t_2) & \dots & \rho(t_1,t_n) \\ \rho(t_2,t_1) & \rho(t_2,t_2) & \dots & \rho(t_2,t_n) \\ \vdots & \vdots & \vdots & \vdots \\ \rho(t_n,t_1) & \rho(t_n,t_2) & \dots & \rho(t_n,t_n) \end{pmatrix},$$

$\boldsymbol{x}$ is the row vector $(x_1, \dots, x_n)$ and $m(\boldsymbol{t})$ is the row vector $(m(t_1), \dots, m(t_n))$.

Problem 2.4. Compute the finite dimensional distributions of the Poisson process N_t.

Hint: first, see Exercise 1.8, it holds

$$\mathbb{P}(N_t = n) = \mathrm{e}^{-\lambda t} \frac{(\lambda t)^n}{n!}, \qquad n \in \mathbb{N}.$$

Next, for any finite sequence of times $t_0 = 0 < t_1 < \dots < t_k$, using the fact that N_t has independent increments, show that

$$\mu_{t_1,\dots,t_k}(n_1, \dots, n_k) = \mathbb{P}(N_{t_1} = n_1, \dots, N_{t_k} = n_k) = \prod_{j=1}^{k} \mathrm{e}^{-\lambda(t_j - t_{j-1})} \frac{(\lambda(t_j - t_{j-1}))^{n_j - n_{j-1}}}{(n_j - n_{j-1})!},$$

for any integers $n_0 = 0 \le n_1 \le \ldots \le n_k$.

Notice that the finite-dimensional distributions of a stochastic process $\{X_t,\ t \in T\}$ always satisfy the following *consistency condition*:

$$\mu_{t_1,\dots,t_m}(A_1 \times \ldots \times A_m) = \mu_{t_1,\dots,t_m,t_{m+1}}(A_1 \times \ldots \times A_m \times E), \tag{2.7}$$

with $t_1 < t_2 < \dots < t_m$ in T.

More interesting is the converse problem: given a family of probability measures $\mu_{t_1,\dots,t_m}$ on $\mathcal{B}(E^m)$, for any choice of m and $t_1 < t_2 < \dots < t_m$ in T, which satisfies (2.7), does there exist a stochastic process (defined on some $(\Omega, \mathcal{F}, \mathbb{P})$) having these as finite dimensional distributions? The answer is affirmative and is given by the Kolmogorov existence theorem. To formulate this we introduce some notation.

Let T be an arbitrary set of indices and E^T be the collection of all functions from T into E; E^T can be thought of as a product space of many copies of E, indexed by T. On this space we consider the *projections* π_t, $\pi_t(x(\cdot)) = x(t)$, $x \in E^T$, $t \in T$. Recall that there exists a natural extension to E^T of the Borel σ-field for vector spaces E^k, that we denote by $\mathcal{E}^T$; this is the smallest σ-field which makes each $\pi_t : E^T \to E$ measurable.

Let us define on $(E^T, \mathcal{E}^T)$ the *coordinate process* $Z = \{Z_t,\ t \in T\}$,

$$Z_t : E^T \to E, \quad Z_t(x) = \pi_t(x) = x(t), \qquad x \in E^T.$$

If we provide the space $(E^T, \mathcal{E}^T)$ with a probability measure $\mathbb{P}$, we obtain that $\{Z_t,\ t \in T\}$ is a stochastic process.

Theorem 2.13. *Let $\{\mu_{t_1,\dots,t_m}\}$ be a family of Borel probability distributions satisfying the consistency condition (2.7). Then there exists a unique probability measure $\mathbb{P}$ on $(E^T, \mathcal{E}^T)$ such that the coordinate process $\{Z_t,\ t \in T\}$ on $(E^T, \mathcal{E}^T, \mathbb{P})$ has finite dimensional distributions which coincide with $\mu_{t_1,\dots,t_m}$, for any choice of $t_1 < t_2 < \dots < t_m$ in T.*

For our pourposes, it is not essential to prove this theorem (the interested reader is referred, for instance, to `Billingsley` [Bi95]). Hence, we only make a remark on the proof in the case $T = \mathbb{N}$.

First one defines $\mathbb{P}$ on the algebra $\mathcal{A}$ of all cylindrical sets

$$I_{t_1,\dots,t_n,A} = \{x \in E^{\mathbb{N}} \, : \, (x_{t_1}, \dots, x_{t_n}) \in A\},$$

where $n \in \mathbb{N}$, $t_1 < \ldots < t_n \in \mathbb{N}$ and $A \in \mathcal{B}(E^n)$, setting

$$\mathbb{P}(I_{t_1,\dots,t_n,A}) := \mu_{t_1,\dots,t_n}(A). \tag{2.8}$$

This definition is meaningful thanks to the consistency condition. It is easy to check that $\mathbb{P}$ is a pre-measure on $\mathcal{A}$. The crucial step of the proof is now to show that $\mathbb{P}$ is continuous (to this purpose, one also uses that Borel measures on E are regular). After the main step is achieved, an application of the Caratheodory extension theorem allows to get that there exists a unique probability measure $\mathbb{P}$ on the σ-algebra generated by $\mathcal{A}$, which coincides with $\mathcal{E}^T$, such that (2.8) holds. Finally, it is clear that finite dimensional distributions of the coordinate process Z_t coincide with the measures $\mu_{t_1,\dots,t_m}$.

Next, we investigate a question arising from the previous result.

Remark 2.14. It must be noticed that in the above construction the choice of E^T is unsatisfactory under several points of view. Let $T = [0, \infty)$ and $E = \mathbb{R}$. It turns out that on the space $\mathbb{R}^{[0,\infty)}$ finite dimensional distributions do not characterize important properties of the trajectories of a process, as for instance continuity.

Indeed, the set $\mathcal{C}$ of continuous functions from $[0, \infty)$ into $\mathbb{R}$ does not belong to $\mathcal{E}^T$. This fact can be explained as follows: given a real function on $[0, \infty)$ it is not possible to establish if it is continuous by looking only to its values on a countable set of points in $[0, \infty)$.

In the next exercise we propose to show a characterization of the σ-algebra $\mathcal{E}^T$, when $T = [0, \infty)$ and $E = \mathbb{R}^n$.

Problem 2.5. Let us consider $\mathcal{E}^T$, when $T = [0, \infty)$ and $E = \mathbb{R}^n$. Show that the family $\mathcal{H}$ of all sets of the form

$$\{x \in E^T \, : \, (x(t_1), x(t_2), \dots, x(t_n), \dots) \in B\},$$

for any choice of sequences of times $t_1 < \ldots < t_n < \ldots$ in T and of sets $B \in \prod_{i \in \mathbb{N}} (\mathcal{E})_i$ (here $\prod_{i \in \mathbb{N}} (\mathcal{E})_i$ denotes the product σ-algebra, where $(\mathcal{E})_i = \mathcal{E}$, $i \in \mathbb{N}$) is a σ-field in E^T.

Deduce that $A \in \mathcal{E}^T$ if and only if there is a sequence of times $t_1 < \ldots < t_n < \ldots$ in T and a set $B \in \prod_{i \in \mathbb{N}} (\mathcal{E})_i$ so that $A = \{x \in E^T \, : \, (x(t_1), x(t_2), \dots, x(t_n), \dots) \in B\}$.

Example 2.15. Let us show that it is possible to construct a process with the same finite-dimensional distributions of the Poisson process $\{N_t,\ t \geq 0\}$, but which exhibits completely different properties of the paths (for instance, they are neither right-continuous, nor monotone non-decreasing).

Take the function $f(t) = t$ if t is rational, $f(t) = 0$ otherwise. Let X_1 be the first waiting time in the definition of the Poisson process $\{N_t\}$ and consider the process $M_t = N_t + f(t + X_1)$. We have

$\mathbb{P}(f(t+X_1)=0)=1$, since the complementary event is $X_1 \in \mathbb{Q}-t$ ($\mathbb{Q}$ is the set of rational point), and the probability that a continuous random variable takes value in a countable set is zero. Thus

$$\mathbb{P}(M_t = N_t) = 1 \qquad \text{for all } t \geq 0,$$

and the process M_t has the same finite-dimensional distributions of N_t, but its paths are everywhere discontinuous, are not integer valued nor monotone.

2.3.1 A construction of Brownian motion using Kolmogorov's theorem

Since a real Brownian motion is a Gaussian process, its finite dimensional distributions are specified by the mean function $m(t)=0$ and covariance function $\rho(t,s)=\min\{s,t\}$ as in (2.6): for $0<t_1<t_2<\cdots<t_n$

$$\begin{aligned}\mu_{t_1,\ldots,t_n}&(A_1\times\cdots\times A_n)\\ &=\frac{1}{(2\pi)^{n/2}\sqrt{\det\Sigma}}\int\limits_{A_1\times\cdots\times A_n}\exp\Big(-\frac{1}{2}\langle\boldsymbol{x},\Sigma^{-1}\boldsymbol{x}\rangle\Big)\,\mathrm{d}x_1\ldots\mathrm{d}x_n,\end{aligned} \tag{2.9}$$

where Σ is the $n\times n$ matrix

$$\Sigma = \begin{pmatrix} t_1 & t_1 & \ldots & t_1\\ t_1 & t_2 & \ldots & t_2\\ \vdots & \vdots & \vdots & \vdots\\ t_1 & t_2 & \ldots & t_n\end{pmatrix}$$

$\det\Sigma = t_1\cdot(t_2-t_1)\ldots(t_n-t_{n-1})>0$ and $\boldsymbol{x}$ is the row vector $(x_1,\ldots,x_n)$. Consistency in this case is obvious. Thus Kolmogorov's theorem 2.13 applies: there exists a stochastic process $\{B_t,\ t\geq 0\}$ corresponding to the finite dimensional distributions $\mu_{t_1,\ldots,t_n}$ (see (2.9)) on the probability space $(\mathbb{R}^{[0,+\infty)},\mathcal{E}^{[0,+\infty)},\mathbb{P})$ with a suitable probability measure $\mathbb{P}$.

Note that a similar construction can be applied, more generally, for proving the existence of a Brownian motion with values in $\mathbb{R}^d$, $d\geq 1$.

However, the Kolmogorov construction does not provide continuity of trajectories of Brownian motion. Actually, we have already observed that the set of continuous functions is not contained in $\mathcal{E}^T$. A second theorem, again due to Kolmogorov, allows to overcome this difficulty.

2.4 Continuity of Brownian motion paths

Recall that a function $f: E\to\mathbb{R}$ is called α-Hölder continuous on $D\subset E$, $\alpha\in(0,1)$, if there exists a constant $c>0$ such that

$$|f(x)-f(y)|\leq c|x-y|^\alpha \qquad \text{for all } x,y\in D.$$

Let $\{X_t,\ t\in T\}$ be a stochastic process (but the result holds, more generally, for a random field indexed by $T\subset\mathbb{R}^d$, $d\geq 2$); the following result, due to Kolmogorov, asserts that a certain degree of regularity in the mean implies some regularity (Hölder continuity) of the trajectories.

Theorem 2.16. *Let $X = \{X_t, t \in T\}$ be a stochastic process in an Euclidean space $(E, \mathcal{E})$ and assume there exist positive constants α, β and c such that*

$$\mathbb{E}\left(|X_t - X_s|^\beta\right) \le c|t-s|^{1+\alpha}. \tag{2.10}$$

Then there exists a version Y of X with continuous sample paths. Further, for any $\gamma < \frac{\alpha}{\beta}$ trajectories $t \mapsto Y_t(\omega)$ are γ-Hölder continuous on every bounded interval, for any $\omega \in \Omega$.

We give the proof in the particular case of $T = (0,1)$ for simplicity of notation. The proof uses the following purely deterministic lemma.

Lemma 2.17. *Let D be the set of all dyadic numbers in (0,1), i.e., $D = \{\frac{k}{2^n} \ : \ n \in \mathbb{N},\ k = 1, \dots, 2^n - 1\}$. Let $f : D \to \mathbb{R}$ be a function and let $\alpha \in (0,1]$. Consider the quantity*

$$c_n = \sup\left\{\frac{|f(x) - f(y)|}{1/2^{\alpha n}} \ : \ x,\, y \in D,\ |x-y| = 1/2^n\right\}, \quad n \in \mathbb{N}.$$

Then $\sup_{n \in \mathbb{N}} c_n < +\infty$ if and only if f is α-Hölder on D.

Proof (Theorem 2.16). Assume $T =]0,1[$ and denote by D the set of dyadic numbers in T; also, denote by $D_n \subset D$ the set of points with coordinates $k/2^n$, for $k = 1, \dots, 2^n - 1$. Finally, fix some $\gamma < \frac{\alpha}{\beta}$.

Step 1. Set

$$B_n = \{\omega \, : \, |X_i - X_j| > 2^{-n\gamma} \text{ for some } i,j \in D_n \text{ with } |i-j| = 1/2^n\}.$$

Then, using (2.10) and Chebyshev inequality, we have, for $i, j \in D_n$ such that $|i-j| = 1/2^n$,

$$\mathbb{P}(|X_i - X_j| > 2^{-n\gamma}) \le 2^{n\gamma\beta}\, \mathbb{E}[|X_i - X_j|^\beta] \le 2^{n\gamma\beta}\, 2^{-n(\alpha+1)} = 2^{n(\gamma\beta - \alpha - 1)}.$$

This implies that

$$\mathbb{P}(B_n) \le 2 2^n\, 2^{n(\gamma\beta-\alpha-1)} = C 2^{-(\alpha-\beta\gamma)n}, \ \ n \in \mathbb{N}, \tag{2.11}$$

for some $C > 0$.

Step 2. Consider $B = \limsup_n B_n = \bigcap_{n=1}^\infty \bigcup_{k=n}^\infty B_k$; by (2.11) we know that $\sum_n \mathbb{P}(B_n)$ converges, hence $\mathbb{P}(B) = 0$ by the first Borel-Cantelli lemma. Hence $\mathbb{P}(\Omega \setminus B) = 1$ and for any $\omega \in \Omega \setminus B$ there exists $\nu = \nu(\omega)$ such that, for any $n \ge \nu$ we have

$$|X_i(\omega) - X_j(\omega)| \le \frac{1}{2^{n\gamma}}, \quad \text{for any } i,j \in D_n,\ |i-j| = 1/2^n.$$

It follows that, for any $\omega \notin B$, there exists $M = M(\omega) > 0$ such that, for any $n \ge 1$,

$$|X_i(\omega) - X_j(\omega)| \le M(\omega)\, \frac{1}{2^{n\gamma}}, \qquad \text{for any } i,j \in D_n,\ |i-j| = 1/2^n.$$

Step 3. By Lemma 2.17, for every $\omega \notin B$ the trajectory $i \mapsto X_i(\omega)$ is $\gamma-$Hölder continuous as i varies in the set D of dyadic points of T, i.e., there exists $C = C(\omega) > 0$ such that, if $i, j \in D$ then

$$|X_i(\omega) - X_j(\omega)| \le C(\omega)|i-j|^\gamma.$$

In particular, the trajectory is uniformly continuous on D.

Step 4. For every $\omega \notin B$ there exists a unique extension of $X_i(\omega)$, $i \in D$, to a trajectory $\tilde{X}_t(\omega)$ defined for $t \in T$. This extension is Hölder continuous of exponent γ.

Step 5. It remains to prove that $\tilde{X}_t = X_t$ with probability 1, $t \in T$. This clearly holds when $t \in D$; now take $t \in T$ and a sequence of dyadic rationals (t_n) such that $t_n \to t$. Using again (2.10), we show that $X_{t_n} \to X_t$ in probability; hence, passing if necessary to a subsequence, $X_{t_n} \to X_t$ almost surely. But since $X_{t_n} = \tilde{X}_{t_n}$, almost surely, it is $\tilde{X}_t = X_t$ almost surely, $t \in T$, which concludes the proof. □

Let now $B = \{B_t,\ t \geq 0\}$ be a Brownian motion. Recall, from Remark 2.7 that for every $n \geq 1$

$$\mathbb{E}[|B_t - B_s|^{2n}] \leq C_n (t-s)^n.$$

By applying Kolmogorov's continuity theorem 2.16 we have that trajectories of the Brownian motion are continuous, and Hölder continuous of exponent γ for every $\gamma < \frac{n-1}{2n}$, i.e., due to the arbitrarity of n, for every $\gamma < \frac{1}{2}$.

Exercises

Problem 2.6. Provide a complete proof of Lemma 2.17.

Problem 2.7. Given a real Brownian motion $\{B_t,\ t \geq 0\}$ on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$, we can construct several other processes which share the same property.

1. (*Omogeneity*) For any fixed $s \geq 0$, the process $X_t = B_{t+s} - B_s$ is a Brownian motion with respect to the filtration $\{\mathcal{F}_{t+s},\ t \geq 0\}$.
2. (*Simmetry*) The process $Y_t = -B_t$ is a Brownian motion with respect to the initial filtration $\{\mathcal{F}_t\}$.
3. (*Scale change*) Given a real number $c > 0$, the process $W_t = cB_{t/c^2}$ is a Brownian motion with respect to the filtration $\{\mathcal{F}_{t/c^2},\ t \geq 0\}$.

Next proposition is somewhat surprising; we can phrase it by saying that if we let time flow backward, we still have a Brownian motion.

Proposition 2.18. *Given a real Brownian motion $\{B_t,\ t \geq 0\}$ on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$. The stochastic process $Y = \{Y_t,\ t \geq 0\}$ defined by*

$$Y_t = \begin{cases} tB_{1/t} & t > 0 \\ 0 & t = 0 \end{cases}$$

is again a Brownian motion with respect to the completed natural filtration $\mathcal{F}^B = \{\mathcal{F}_t^B,\ t \geq 0\}$.

Proof. We verify that Proposition 2.11 applies. Property a) is obvious; property b) follows easily since Y is again a Gaussian family. Indeed, for any choice of times $0 \leq t_1 < t_2 < \cdots < t_m$, we can write $(Y_{t_1}, \ldots, Y_{t_m})$ as a linear transformation of $(B_{t_1}, \ldots, B_{t_m})$ and then apply Proposition 2.2.

Finally, if $s \leq t$:

$$\mathbb{E}[Y_t Y_s] = st\mathbb{E}[B_{1/t}B_{1/s}] = st\frac{1}{s} \wedge \frac{1}{t} = s = s \wedge t$$

and property c) follows.

Thus, it remains to prove that sample paths are almost surely continuous. This clearly holds for $t > 0$, i.e., there exists a $\mathbb{P}$-null (or negligible) set N^* such that, for every $\omega \in \Omega \setminus N^*$, the mapping $t \mapsto Y_t(\omega)$ is continuous on $(0, +\infty)$. We shall now show that there exists a negligible set N such that outside N, the mapping $t \mapsto Y_t(\omega)$ is continuous on $[0, +\infty)$.

Notice that, since $Y_t - Y_s$ has Gaussian law $\mathcal{N}(0, t-s)$,

$$\mathbb{E}[|Y_t - Y_s|^4] = 3(t-s)^2.$$

Thanks to Kolmogorov's theorem 2.16, there exists a modification $\tilde{Y}$ of Y with continuous paths; hence, for every $t \geq 0$, there exists a $\mathbb{P}$-null set N_t such that

$$Y_t(\omega) = \tilde{Y}_t(\omega) \qquad \forall \omega \in \Omega \setminus N_t.$$

Let N be the union of N^* and the sets N_τ, as τ varies in the positive rational numbers; N is a countable union of negligible sets, hence its probability is zero. For every $\omega \in \Omega \setminus N$, trajectories $t \mapsto Y_t(\omega)$ and $t \mapsto \tilde{Y}_t(\omega)$ are continuous functions on $]0, \infty[$ which coincide on the rational numbers, hence they must coincide on the whole half line $[0, \infty[$. But since $\tilde{Y}_t(\omega)$ is continuous up to 0, so it is $Y_t(\omega)$, and the proof is complete.

□

Addendum. The Borel-Cantelli lemmas

For the sake of completeness, we recall here the Borel-Cantelli lemma. Actually, it consists of two parts, sometimes called the first (second) Borel-Cantelli lemma. Both results concern the limit behaviour of a sequence of events, and as such they are often useful instruments in proving almost sure properties.

Lemma 2.19 (Borel-Cantelli).

(a) *Assume that* $\{A_n,\ n \in \mathbb{N}\}$ *is a sequence of events such that*

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) < +\infty.$$

Then

$$\mathbb{P}(\limsup_{n \in \mathbb{N}} A_n) = \mathbb{P}(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k) = \mathbb{P}(\{\omega \colon \omega \in A_n \textit{ for infinite indices } n\}) = 0.$$

(b) *Assume that* $\{A_n,\ n \in \mathbb{N}\}$ *is a sequence of pairwise independent events such that*

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) = +\infty.$$

Then

$$\mathbb{P}(\limsup_{n \in \mathbb{N}} A_n) = 1.$$