Note 10

Construction of general stochastic processes. The Daniell-Kolmogorov theorem.

1 Introduction

In this brief note we review the general construction of a stochastic process via its marginal distributions.

For modelisation purposes we would like to consider families of random variable $(X_t)_{t\in I}$ indexed by a general parameter set I. When I is discrete and finite then there is no problem to see the family as a single sample of a vector valued random variable with values in the product measure space $(E^I,\mathscr{E}^I)$ where $(E,\mathscr{E})$ is the measure space where each X_t is defined, i.e. X_t: $(\Omega,\mathscr{F})\to(E,\mathscr{E})$ for all $t\in I$. Taking π_t: $(E^I,\mathscr{E}^I)\to(E,\mathscr{E})$ to be the measurable map giving the coordinate projecton on the factor t we can simply define $X:\Omega\to(E^I,\mathscr{E}^I)$ with $(X(\omega))_t=X_t(\omega)$ for all $t\in I$ and $\omega\in\Omega$ and obtain that $\pi_t X=X_t$. In this way we have replaced the family of random variables $(X_t)_t$ each of them taking values on E with one single random variable X taking values on $(E^I,\mathscr{E}^I)$ where $\mathscr{E}^I$ is the product σ-algebra, i.e. the σ-algebra generated by sets of the form $\prod_{i\in I}A_i\subseteq E^I$ where $A_i\in\mathscr{E}$ for all $i\in I$.

As soon as I is not finite or not even countable this procedure has to be clarified. As an example we would like to allow I to be $\mathbb{N}$, $\mathbb{Z}$ or even $\mathbb{R}$ or $\mathbb{R}^n$, or even larger spaces like the space of Schwartz test functions $\mathscr{S}(\mathbb{R}^d)$ on some Euclidean space $\mathbb{R}^d$. In the case the parameter set is not one dimensional, we usually speak of random fields instead of random processes. Anyway in the following the explicit form of I does not play particular role, and this is quite remarkable and due to the fact that the axioms of a σ-algebra already put very stringent constraints on the objects one can construct that already the example of $I=\mathbb{N}$ or $I=\mathbb{Z}$ essentially saturates those constraints in such a way that going from $I=\mathbb{N}$ to $I=\mathscr{S}(\mathbb{R}^d)$ does not add additional difficulties.

We are looking at a definition for $\mathscr{E}^I$ which retains the property that for all $t\in I$ the maps π_t: $(E^I,\mathscr{E}^I)\to(E,\mathscr{E})$ remains measurable. Therefore we can just pose this as a definition and let $\mathscr{E}^I$ be the smallest σ-algebra which makes all the π_t: $E^I\to(E,\mathscr{E})$ measurable. The family of sets of the form

$$\pi_J^{-1}(A),\qquad J\subseteq I, J \text{ finite}, A\in\mathscr{E}^J,$$

where $\pi_J: E^I\to E^J$ is the projection on the coordinates in J constitute an algebra in $\mathscr{P}(E^I)$. Sets of the form $\pi_J^{-1}(A)$ are called cylinder sets. The family of *special cylinder sets* of the form

$$\pi_J^{-1}\left(\prod_{t\in J}A_t\right), J\subseteq I, J \text{ finite}, A_t\in\mathscr{E} \text{ for all } t\in J,$$

is a π-system inside the algebra of cylinder sets. Both systems generate $\mathscr{E}^I$: any σ-algebra which contains the special cylinder sets makes every projection π_t measurable and the smallest is indeed $\mathscr{E}^I$.

The definition of the product σ-algebra is conceived to achieve the following equivalence.

Lemma 1. *A map $X: (\Omega, \mathscr{F}) \rightarrow (E^I, \mathscr{E}^I)$ is measurable iff $\pi_t X: (\Omega, \mathscr{F}) \rightarrow (E, \mathscr{E})$ is measurable for all $t \in I$.*

Proof. Note that $X: (\Omega, \mathscr{F}) \rightarrow (E^I, \mathscr{E}^I)$ is measurable iff $X^{-1}(A) \in \mathscr{F}$ for all A which generates $\mathscr{E}^I$ but this means that we can check the condition for the special cylinders $A = \pi_J^{-1}(\prod_{t \in J} B_t)$ for $B_t \in \mathscr{E}$ and $J \subseteq I$ finite and this is indeed implied by the measurability of each of the projections $(\pi_t)_{t \in J}$ since $\pi_J^{-1}(\prod_{t \in J} B_t) = \cap_{t \in J} \pi_t^{-1}(B_t) \in \mathscr{F}$. The opposite implication derives immediately from the composition of measurable functions since π_t is measurable from $(E^I, \mathscr{E}^I)$ to $(E, \mathscr{E})$ for all $t \in I$. □

The interest of this result is that, as soon as I is not finite, $\mathscr{E}^I$ contains very interesting events which are not simply determined by any of the finite dimensional projections, think for example to some tail events.

Given a process $X: (\Omega, \mathscr{F}) \rightarrow (E^I, \mathscr{E}^I)$ we call the measure μ_X on $(E^I, \mathscr{E}^I)$ given by $\mu_X(A) = \mathbb{P}(X \in A)$ for all $A \in \mathscr{E}^I$ its law. Moreover as already seen, on the space $(E^I, \mathscr{E}^I)$ we can always realize the stochastic process $(X_t: \Omega \rightarrow E)_{t \in I}$ by taking $X_t(\omega) = \omega_t$ so that $X(\omega) = \omega$. This is called the canonical process.

2 Daniell-Kolmogorov's theorem

Another advantage of working with the σ-algebra $\mathscr{E}^I$ is that the law of a stochastic process is characterized by its finite dimensional marginals. Indeed the special cylinders form a π-system which generates $\mathscr{E}^I$ and therefore two probabilities on $(E^I, \mathscr{E}^I)$ which coincide on all special cylinders must also coincide on $\mathscr{E}^I$ by the standard $\pi - \lambda$ argument (consider the Λ-system of sets on which the two measures coincide…).

The following theorem, due to Kolmogorov (and independently to Daniell) state the existence of a probability on $(E^I, \mathscr{E}^I)$ is ensured by the existence of a family of consitent probabilities for the special cylinders.

We say that a family of probabilites $(\mu_J)_J$ indexed by the finite subsets of I is consistent if for any $J' \subseteq J$ with J finite and with $\pi_{J,J'}: E^J \rightarrow E^{J'}$ the canonical projection from E^J to $E^{J'}$we have $\mu_J \circ \pi_{J,J'}^{-1} = \mu_{J'}$.

Theorem 2. (Daniell-Kolmogorov) *Assume that $(\mu_J \in \Pi(E^J, \mathscr{E}^J))_{J \subseteq I, J \text{ finite}}$ is a consistent family of probabilites which are inner regular. Then there exist a unique measure μ on $(E^I, \mathscr{E}^I)$ such that $\mu \circ \pi_J^{-1} = \mu_J$ for all finite $J \subseteq I$.*

Proof. On the algebra of cylinder sets we can define an additive set function μ_0 as follows. Let A be a cylinder set, then $A = \pi_J^{-1}(A_J)$ for some $J \subseteq I$ finite and $A_J \in \mathscr{E}^J$. Then let

$$\mu_0(A) = \mu^J(A_J).$$

The consistency of the family is needed to ensure that this definition is well posed, indeed if we have $A = \pi_{J'}^{-1}(A_{J'})$ for another $J' \subseteq I$ finite, we also have $A = \pi_{J'}^{-1}(A_{J'}) \cap \pi_J^{-1}(A_J) = \pi_{J \cup J'}^{-1}(A_{J \cup J'})$ for $A_{J \cup J'} \in \mathscr{E}^{J \cup J'}$ with $A_{J \cup J'} = \pi_{J \cup J', J}^{-1}(A_J) = \pi_{J \cup J', J'}^{-1}(A_{J'})$, then by consistency

$$\mu^J(A_J) = \mu^{J \cup J'}(\pi_{J \cup J', J}^{-1}(A_J)) = \mu^{J \cup J'}(A_{J \cup J'}) = \mu^{J \cup J'}(\pi_{J \cup J', J'}^{-1}(A_{J'})) = \mu^{J'}(A_{J'})$$

and therefore the definition does not depends on the particular choice of the "base" of the representation of A as a cylinder set. It is clear that $\mu_0(E^I) = 1$ and that μ_0 is additive, since any pair of cylinders can be represent as preimages of the same projection. Less clear is the σ-additivity which is where we need the additional assumption that E is compact and metrizable, but which is equivalent to prove that for a family $(A_n)_n$ of cylinder sets so that $A_n \supseteq A_{n+1}$ and $A_n \downarrow \emptyset$ we have $\mu_0(A_n) \to 0$. We are going to prove that this is necessarily true by showing that if $\mu_0(A_n) \geqslant 2\varepsilon$ for all n then $\cap_n A_n \neq \emptyset$. Since each A_n is a cylinder set, there exists $J_n \subseteq I$ finite and $A_{J_n} \in \mathscr{E}^{J_n}$ such that $A_n = \pi_{J_n}^{-1}(A_{J_n})$. Since μ^{J_n} is inner regular there exists a compact set $K_n \subseteq A_n$ such that $\mu^{J_n}(K_n) \geqslant \mu^{J_n}(A_n) - 2^{-n}\varepsilon$. Then if we let $H_n = \pi_{J_n}^{-1}(K_n)$ we have

$$\mu_0(\cap_{k=1}^n H_k) \geqslant \mu_0(\cap_{k=1}^n A_k) - \sum_{k=1}^n \mu_0(A_k \backslash H_k) \geqslant 2\varepsilon - \sum_{k=1}^n \varepsilon 2^{-k} \geqslant \varepsilon$$

and in particular for any n, $\cap_{k=1}^n H_k \neq \emptyset$. Let $x_n \in \cap_{k=1}^n H_k$, that is $\pi_{J_k} x_n \in \cap_{\ell=1}^k K_\ell$ for all $k \leqslant n$. By compactness there exists a subsequence $(\pi_{J_k} x_{n_r})_r$ such that $\lim_r \pi_{J_k} x_{n_r} \in \cap_{\ell=1}^k K_\ell$. By a diagonal argument we can extract a further subsequence such that $\lim_r \pi_{J_k} x_{n_r} = x^k \in \cap_{\ell=1}^k K_\ell$ for all $k \geqslant 0$. Clearly $\pi_{J_m} x^k = x^m$ for all $k \geqslant m$. Therefore there exists $x \in E^I$ such that $\pi_{J_k} x = x^k \in \cap_{\ell=1}^k K_\ell$ for all k and as a consequence $x \in \cap_{\ell=1}^\infty K_\ell$. But this implies that $\cap_{\ell=1}^\infty K_\ell$ is not empty. By contradiction we established that μ_0 is σ-additive and therefore that by Caratheodory extension theorem can be extended to a probability measure on $(E^I, \mathscr{E}^I)$. Uniqueness comes from the π-λ argument sketched above. □

In particular the assumption of inner regularity holds if E is a Polish space as the following lemma establish.

Lemma 3. *Any probability on a Polish space $(\Omega, \mathscr{F})$ endowed with its Borel σ-algebra is inner regular.*

Proof. Let $B_r(x)$ denote the (closed) ball around $x \in \Omega$ of radius r. Since Ω is separable, for any n there exists a sequence of points $(x_k^n)_k$ such that $\Omega = \cup_k B_{1/n}(x_k^n)$ and by σ-additivity there exists an l_n such that

$$\mathbb{P}(\cup_{k=1}^{l_n} B_{1/n}(x_k^n)) \geqslant 1 - \varepsilon 2^{-n}.$$

Then the set $K = \cap_n \cup_{k=1}^{l_n} B_{1/n}(x_k^n)$ is closed and

$$\mathbb{P}(K^c) = \mathbb{P}(\cup_n (\cup_{k=1}^{l_n} B_{1/n}(x_k^n))^c) \leqslant \sum_n \varepsilon 2^{-n} \leqslant \varepsilon.$$

Moreover $K \subseteq \cup_{k=1}^{l_n} B_{1/n}(x_k^n)$ for every n, therefore K is closed and totally bounded and therefore compact since Ω is complete. So Ω can be approximated by a compact set with arbitrary small loss of measure. From this one can also deduce that any closed set has this property and from Dynkin's theorem that all the measurable set have the same property. □

3 Gaussian processes

To show the generality of the above construction we start by considering a real Hilbert space H with scalar product $\langle \cdot, \cdot \rangle$. To each finite subset $J \subseteq H$ we associate the measure μ^J on $\mathbb{R}^J$ given by the law $\mathscr{N}_{\mathbb{R}^J}(0, C^J)$ of the Gaussian vector $(X_h)_{h \in J}$ with mean zero and covariance

$$C_{h,h'}^J = \langle h, h' \rangle, \qquad h, h' \in J.$$

Note that his is possible since the covariance matrix $(C^J_{h,h'})_{h,h'}$ is positive semidefinite, since it is so the scalar product of H:

$$\sum_{h,h'\in J} \lambda_h \lambda_{h'} C^J_{h,h'} = \sum_{h,h'\in J} \lambda_h \lambda_{h'} \langle h, h' \rangle = \left\| \sum_{h\in J} \lambda_h h \right\|_H^2 \geqslant 0$$

for all choice of $(\lambda_h \in \mathbb{R})_{h\in J}$. An alternative characterisation of the measure $\mathcal{N}_{\mathbb{R}^J}(0, \mathrm{C}^J)$ is given by its caracteristic function

$$\int_{\mathbb{R}^J} e^{i(\sum_{h\in J}\lambda_h x_h)} \mu^J(\mathrm{d}x) = e^{-\frac{1}{2}\sum_{h,h'\in J}\lambda_h\lambda_{h'}C^J_{h,h'}}.$$

An easy computation shows that, if $J \subseteq J'$ then $\mu^J \circ \pi_{J,J'} = \mu^{J'}$, for example from the caracteristic function, since

$$\int_{\mathbb{R}^{J'}} e^{i(\sum_{h\in J'}\lambda_h x_h)} (\mu^J \circ \pi_{J,J'})(\mathrm{d}x) = \int_{\mathbb{R}^J} e^{i(\sum_{h\in J'}\lambda_h(\pi_{J,J'}x)_h)} \mu^J(\mathrm{d}x)$$

$$= \int_{\mathbb{R}^J} e^{i(\sum_{h\in J}\hat{\lambda}_h x_h)} \mu^J(\mathrm{d}x) = e^{-\frac{1}{2}\sum_{h,h'\in J}\hat{\lambda}_h\hat{\lambda}_{h'}C^J_{h,h'}} = e^{-\frac{1}{2}\sum_{h,h'\in J'}\lambda_h\lambda_{h'}C^{J'}_{h,h'}} = \int_{\mathbb{R}^{J'}} e^{i(\sum_{h\in J'}\lambda_h x_h)} \mu^{J'}(\mathrm{d}x)$$

where $\hat{\lambda}_h = \mathbb{1}_{h\in J'}\lambda_h$. Therefore the family $(\mu^J)_J$ is consistent and by the Daniell-Kolmogorov theorem it defines a measure $\mathbb{P}$ on the measure space $(\mathbb{R}^H, \mathscr{B}(\mathbb{R})^{\otimes H})$ under which the canonical process $(X_h)_{h\in H}$ has Gaussian finite dimensional marginals with covariance given by the scalar product of H. In particular we have

$$\mathbb{E}e^{i\lambda X_h} = e^{-\frac{1}{2}\lambda^2\|h\|^2}$$

and this suffices to characterise $\mathbb{P}$ since for any finite J, choosing $h = \sum_{h\in J} \lambda_h h_h$ we have

$$\mathbb{E}e^{iX_h} = e^{-\frac{1}{2}\|h\|^2} = e^{-\frac{1}{2}\sum_{h,h'\in J}\lambda_h\lambda_{h'}\langle h,h'\rangle}$$

which shows that the law of the vector $(X_h)_{h\in J}$ is precisely μ^J for all finite J. Note that

$$\mathbb{E}(\lambda X_h - X_{\lambda h})^2 = \lambda^2\|h\|^2 + \|\lambda h\|^2 - 2\lambda\langle h, \lambda h\rangle = 0$$

therefore, almost surely $\lambda X_h = X_{\lambda h}$. Similarly one can prove that $X_{h_1+h_2} = X_{h_1} + X_{h_2}$ almost surely for any $h_1, h_2 \in H$. The map $h \in H \mapsto X_h \in L^2(\mathbb{P})$ is an isometry.