PROBABILITY 2018
HANDOUT 1: PROBABILITY SPACES

GIOVANNI PISTONE

Contents

1. The σ-algebra generated by a finite partition, the Borel σ-algebra

The algebra of simple functions; the algebra generated by a finite partition; topology of $\mathbb{R}^n$; Borel σ-algebra.

2. Probability measure on a partition. Probability measure on a denumerable set

Elementary examples. Bernoulli trials.

3. Dynkin lemma

1. *Let $\mathcal{I}$ be a π-system. The family of all real functions of the form $\alpha_0 + \sum_{j=1}^n \alpha_j \mathbf{1}_{I_j}$, $n \in \mathbb{N}$, $\alpha_j \in \mathbb{R}$, $j = 0, \dots, n$, $I_j \in \mathcal{I}$ is a vector space and it is stable for multiplication.*

It is a vector space by construction. Stability under multiplication follows from $\mathbf{1}_I \mathbf{1}_J = \mathbf{1}_{I \cap J}$.

2. *A family $\mathcal{H}$ of subsets of S is a σ-algebra if, and only if, it is both a d-system and a π-system.*

If $\mathcal{H}$ is a σ-algebra, then it is a π-system and a d-system. A d-system contains the total set S, is stable for the complement and stable for increasing unions. We need only to show it is closed under intersection, which is true because it is a π-system.

3. *If $\mathcal{I}$ is a π-system, then $d(\mathcal{I}) = \sigma(\mathcal{I})$.*

Note that $d(\mathcal{I}) \subset \sigma(\mathcal{I})$. We want to show that $d(\mathcal{I})$ is a π-system, because in such a case is a σ-algebra containing $\mathcal{I}$, hence the equality. See D. Williams p. 194.

4. *Any d-system $\mathcal{D}$ that contains a π-system $\mathcal{I}$ contains the σ-algebra generated by the π-system.*

In fact, $\mathcal{D} \supset d(\mathcal{I}) = \sigma(\mathcal{I})$.

5. *If two probability measures on the same measurable space $(S, \mathcal{S})$ agree on a π-system $\mathcal{I}$ they are equal on $\sigma(\mathcal{I})$.*

Date: Version February 16, 2018.

Let $\mathcal{D} = \{A \in \mathcal{S} | \mu_1(A) = \mu_2(A)\}$. Then $\mathcal{D}$ is a d-system that contains $\mathcal{I}$.

(1) A probability on a denumerable set is characterized by its value on points.
(2) A probability of $\mathbb{R}$ is characterized by its distribution function or by its survival function.
(3) A probability of $\mathbb{R}^n$ is characterized by its value on boxes.
(4) A probability on $\{0,1\}^{\mathbb{N}}$ is characterized by its value on cylinder sets.
(5) The distribution of a stochastic process is characterized by its finite dimensional distributions.

4. Monotone class theorems

6. *Let $\mathcal{H}$ be a vector space of real functions of a set S and assume $\mathbf{1} \in \mathcal{H}$. Assume moreover that*

(1) *$\mathcal{H}$ is a* monotone class *i.e., if for increasing sequence $(f_n)_n \in \mathbb{N}$ of non-negative functions in $\mathcal{H}$ bounded by some $h \in \mathcal{H}$, the function $\vee_n f_n$ belongs to $\mathcal{H}$.*
(2) *$\mathcal{H}$ contains the indicator functions of a π-system $\mathcal{I}$.*

Then, $\mathcal{H}$ contains all measurable functions of $(S, \sigma(I))$ which are bounded by an element of $\mathcal{H}$.

- Let us call $\mathcal{D}$ the set of $A \subset S$ such that $\mathbf{1}_A \in \mathcal{H}$. As $\mathbf{1} \in \mathcal{H}$, then $S \in \mathcal{D}$. If $A, B \in \mathcal{D}$ with $B \subset A$, then $\mathbf{1}_{A \backslash B} = \mathbf{1}_A - \mathbf{1}_B \in \mathcal{H}$, hence $A \backslash B \in \mathcal{D}$. If $(A_n)_n$ is an incresing sequence in $\mathcal{D}$, then $0 \leqslant \mathbf{1}_{A_n} \uparrow \mathbf{1}_{\cup_n A_n}$ and $\cup_n A_n \in \mathcal{D}$. As, by assumption, $\mathcal{D}$ contains a π-system $\mathcal{I}$, then $\mathcal{D}$ is contains $\sigma(\mathcal{I})$ because of ¶4.
- Let f be a non-negative function which is $\sigma(\mathcal{I})$-measurable and bounded by $h \in \mathcal{H}$. There exists a sequence of non-negative simple functions f_n of $\sigma(\mathcal{I})$ such that $f_n \uparrow f$. Each f_n is a linear combination of elements of $\mathcal{H}$, hence it belongs to $\mathcal{H}$ and is bounded by $h \in \mathcal{H}$. It follows that $f \in \mathcal{H}$. Finally, a general measurable f with $|f| \leqslant h \in \mathcal{H}$ is of the form $f = f_+ - f_-$ with $f_+, f_- \leqslant h$.

7. *Let $\mathcal{H}$ be a vector space of real functions of a set S and assume $\mathbf{1} \in \mathcal{H}$. Assume moreover that*

(1) *$\mathcal{H}$ is stable under uniform convergence.*
(2) *$\mathcal{H}$ is a* monotone class *i.e., if for increasing sequence $(f_n)_n \in \mathbb{N}$ of non-negative functions in $\mathcal{H}$ bounded by some $h \in \mathcal{H}$, the function $\vee_n f_n$ belongs to $\mathcal{H}$.*
(3) *$\mathcal{H}$ contains a set $\mathcal{C}$ of bounded functions which is stable for the product.*

Then, $\mathcal{H}$ contains all measurable functions of $(S, \sigma(\mathcal{C}))$ which are bounded by an element of $\mathcal{H}$.

- Let C' be the algebra generated by C and $\mathbf{1}$. It follows that $C' \subset \mathcal{H}$. Choose a maximal element $\mathcal{A}$ of all algebras of bounded functions that contain $\mathcal{C}$ and are contained in $\mathcal{H}$.
- Let us show that $\mathcal{A}$ is stable under bounded uniform convergence. Let $(f_n)_n$ be a bounded sequence in $\mathcal{A}$ and assume f is the uniform limit of (f_n). Then, $f \in \mathcal{H}$ and f is uniformly bounded. The set $\mathcal{A}'$ of all such limits contains $\mathcal{A}$, is contained in $\mathcal{H}$, and it is an algebra. But $\mathcal{A}$ is maximal, then $\mathcal{A}' = \mathcal{A}$.
- The function $x \mapsto |x|$ is the uniform limit on bounded intervals of a sequence of polynomials. Take for example[1] the sequence of polynomials defined by $u_1(y) = 0$ and
$$u_{n+1}(y) = u_n(y) + \frac{1}{2}(y - u_n^2(y)), \quad n \geqslant 1.$$

[1]J. Dieudonné *Foundations of modern analysis* 1960, p. 130

We have for $0 \leqslant y \leqslant 1$ that $u_n(y) \leqslant \sqrt{y}$. Observe that

$$\sqrt{y} - u_{n+1}(y) = \sqrt{y} - u_n(y) - \frac{1}{2}(y - u_n^2(y)) = (\sqrt{y} - u_n(y)) \left(1 - \frac{1}{2}(\sqrt{y} + u_n(y))\right).$$

If $u_n(x) \leqslant \sqrt{y}$ then $\frac{1}{2}(\sqrt{y} + u_n(y)) \leqslant \sqrt{y} \leqslant 1$, hence $u_{n+1}(u) \leqslant \sqrt{y}$, and the recursion argument gives the desired result. It follows that $u_{n=1}(y) - u_n(y) = \frac{1}{2}(y - u_n^2(y)) \geqslant 0$. The sequence $u_n(y)$ is positive increasing, bounded by $\sqrt{y}$, then the limit $u(y)$ exists and it is such that $u(y) = u(y) + \frac{1}{2}(y - u^2(y))$, that is $u(y) = \sqrt{y}$. The limit is uniform.

- Now let $g \in \mathcal{A}$ ans assume $|g| \leqslant 1$. Then $u_n(g)$ is a sequence in $\mathcal{A}$ that converges uniformly to $\sqrt{g^2} = |g| \in \mathcal{A}$. In general,, if $|g| \leqslant k$, then $|g| = k\,|g/k| = k \lim_n u_n(g^2/k^k)$. It follows that $\mathcal{A}$ is stable for $\wedge$, $\vee$, positive part, negative part.
- The sets $C_{g_1,\dots,g_n,a_1,\dots,a_n} = \{x \in S | g_i(x) > a_i, i = 1, \dots, n\}$ form a π-system which is a generating set of $\sigma(\mathcal{A})$. Each $C_{g,a}$ is the monotone limit of elements of $\mathcal{A}$, then $\mathbf{1}_{C_{g_1,\dots,g_n,a_1,\dots,a_n}} \in \mathcal{H}$ and all the conditions of ¶7 are satified.

8. *Examples of application ¶7*

(1) *The set $\mathcal{C}$ of bounded continuous functions on $\mathbb{R}^n$ is stable for the multiplication and generated the Borel σ-algebra. If μ_1, μ_2 are probability measures on $\mathbb{R}^n$, define $\mathcal{H}$ the set of all bounded measurable functions. If $\int g d\mu_1 = \int g d\mu_2$ for all g bounded continuous, then $\mu_1 = \mu_2$.*

(2) *The same argument holds for the set of continuous functions with bounded support.*

(3) *The existence of functions which are infinitely differentiable and with compaxt support is not obvious. See in Wikipedia the article* Non-analytic smooth function.

(4) *The set of functions on $\mathbb{R}^n$ of the form $\exp(-p(x))$, where p is a non-negative polynomial of degree 2, is stable for the product and generates the Borel σ-algebra. Popular in Machine Learning.*

(5) *The vector space generated by trigonometric functions cosmx, $\sin nx$, is stable for the product and generates the Borel σ-algebra of $\mathbb{R}$. The elements are usually called trigonometric polynomials.*

(6) *The set of complex exponentials $\mathbb{R}^n \ni x \mapsto \exp\left(\sqrt{-1}\langle t, x\rangle\right) \in \mathbb{C}$ is stable for multiplication, but the complex exponentials are complex valued. This applies to characteristic functions.*

(7) *The set of real exponentials $\mathbb{R}^n \ni x \mapsto \exp(\langle t, x\rangle) \in \mathbb{R}$ is stable under multiplication, but the real exponential are unbounded.*

COLLEGIO CARLO ALBERTO
E-mail address: `giovanni.pistone@carloalberto.org`