Probability 2018
1

Giovanni Pistone
www.giannidiorestino.it

Version February 16, 2018

Measurable space

Definition

- A family $\mathcal{B}$ of subsets of S is an algebra on S if it contains $\emptyset$ and S, and it is stable for the complements, finite unions, and finite intersection.
- A family $\mathcal{F}$ of subsets of S is a σ-algebra on S if it is an algebra on S and it is stable for denumerable unions and intersections.
- A measurable space is a couple $(S, \mathcal{F})$, where S is a set and $\mathcal{F}$ is a σ-algebra on S.

- Given the family $\mathcal{C}$ of subsets of S, the σ-algebra generated by $\mathcal{C}$ is $\sigma(\mathcal{C}) = \cap \{\mathcal{A} | \mathcal{C} \subset \mathcal{A}$ and $\mathcal{A}$ is a σ-algebra$\}$.
- Examples: the algebra generated by a finite partition; the Borel σ-algebra of $\mathbb{R}$ is generated by the open intervals, or by the closed intervals, or by the intervals, or by the open sets, or by semi-infinite intervals. See Handout 1.

Measure space

Definition

- A measure μ of the measurable space $(S, \mathcal{F})$ is a mapping $\mu \colon \mathcal{F} \to [0, +\infty]$ such that $\mu(\emptyset) = 0$ and for each sequence $(A_n)_{n\in\mathbb{N}}$ of disjoint elements of $\mathcal{F}$, $\mu(\cup_{n\in\mathbb{N}} A_n) = \sum_{i=1}^{\infty} \mu(A_n)$.
- A measure is finite if $\mu(S) < +\infty$; a measure is σ-finite if there is a sequence $(S_n)_{n\in\mathbb{N}}$ in $\mathcal{F}$ such that $\cup_{n\in\mathbb{N}} S_n = S$ and $\mu(S_n) < +\infty$ for all $n \in \mathbb{N}$.
- A probability measure is a finite measure such that $\mu(S) = 1$; a probability space is the triple $(S, \mathcal{F}, \mu)$, where μ is a probability measure.

- Examples: probability measure on a partition; probability measure on a denumerable set. See Handout 1.
- Equivalently, a probability measure is finitely additive and sequentially continuous at $\emptyset$

Product system aka π-system

Definition

Let S be a set. A π-system on S is a family $\mathcal{I}$ of subsets of S which is stable under finite intersection.

- Examples: the family of all points of a finite set and the empty set; the family of open intervals of $\mathbb{R}$; the famililly of closed intervals of $\mathbb{R}$; the family of cadlàg intervals of $\mathbb{R}$; the family of convex (resp. open convex, closed convex) subsets of $\mathbb{R}^2$; the family of open (resp. closed) set in a topological space.
- If $\mathcal{I}_i$ is a π-system of S_i, $i = 1, \ldots, n$, then $\{\times_{i=1}^n I_i | I_i \in \mathcal{I}_i\}$ is a π-system of $\times_{i=1}^n S_i$.
- The family of all real functions of the form $\alpha_0 + \sum_{j=1}^n \alpha_j \mathbf{1}_{I_i}$, $n \in \mathbb{N}$, $\alpha_j \in \mathbb{R}$, $j = 0, \ldots, n$ is a vector space and it is stable for multiplication.

§1.6 of ; Handout 1.

Dynkin system aka d-system

Definition

Let S be a set. A d-system on S is a family $\mathcal{D}$ of subsets of S such that

1. $S \in \mathcal{D}$
2. If $A, B \in \mathcal{D}$ and $A \subset B$, then $B \setminus A \in \mathcal{D}$. (Notice that $S \setminus A = A^c$)
3. If $(A_n)_{n\in\mathbb{N}}$ is an increasing sequence in $\mathcal{D}$, then $\cup_{n\in\mathbb{N}} \in \mathcal{D}$

- Given probabilities μ_i and $i = 1, 2$ on the measurable space $(S, \mathcal{F})$, the family $\mathcal{D} = \{A \in \mathcal{F} | \mu_1(A) = \mu_2(A)\}$ in a d-system.
- Given measurable spaces $(S_i, \mathcal{F}_i)$, $i = 1, 2$, the product space $(S, \mathcal{F}) = (S_1 \times S_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$, $\mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma \{A_1 \times A_2 | A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2\}$, and $x \in S_1$, the family $\mathcal{D} = \{A \in \mathcal{F}_1 \otimes \mathcal{F}_2 | A \cap \{x\} \times S_2 = \{x\} \times A_x, A_x \in \mathcal{F}_2\}$ is a d-system.

§A1.2 of

Dynkin's lemma

Theorem

1. *A family of subsets of S is a σ-algebra if, and only if, it is both a d-system and a π-system.*
2. *If $\mathcal{I}$ is a π-system, then $d(\mathcal{I}) = \sigma(\mathcal{I})$.*
3. *Any d-system that contains a π-system contains the σ-algebra generated by the π-system.*

Theorem

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

§A1.3 of ; Handout 1.

Probability space

Definition

A *probability space* is a triple $(\Omega, \mathcal{F}, \mathrm{P})$ of a sample space Ω (set of possible worlds), a σ-algebra $\mathcal{F}$ on Ω, a probability measure $\mathrm{P}\colon \mathcal{F} \to [0,1]$. An element $\omega \in \Omega$ is a sample point (world); an element $A \in \mathcal{F}$ is an event; the value $\mathrm{P}(A)$ is the probability of the event A.

- Examples: a finite set, all its subsets, a probability function $p\colon \Omega \to \mathbb{R}_{>0}$ such that $\sum_{\omega\in\Omega} p(\omega) = 1$; $\mathbb{Z}_{\geq}$ with all its subsets, and a probability function $p\colon \mathbb{Z}_{\geq} \to \mathbb{R}_{>0}$ such that $\sum_{k=0}^{\infty} p(k) = 1$; the restriction of a probability space to a sub-σ-algebra; the product of two probability spaces.
- Bernoulli trials. Let $\Omega = \{0,1\}^{\mathbb{N}}$ and let $\mathcal{F}_n = \left\{ A \times \{0,1\} \times \{0,1\} \times \cdots \middle| A \subset \{0,1\}^n \right\}$, $\mathcal{F} = \sigma(\mathcal{F}_n\colon n \in \mathbb{N})$. Given $\theta \in [0,1]$, the function $p_n(x_1 x_2 \cdots x_n \cdots) = \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i}$ uniquely defines probability spaces $(\Omega, \mathcal{F}_n, \mathrm{P}_n)$, $n \in \mathbb{N}$, such that $\mathrm{P}_{n+1}|_{\mathcal{F}_n} = \mathrm{P}_n$, hence a probability measure P on $\mathcal{F}$.

$\limsup$ and $\liminf$

Definition

- Let $(a_n)_{n\in\mathbb{N}}$ be a sequence of real numbers.

$$\limsup_{n\to\infty} a_n = \bigwedge_{m\in\mathbb{N}} \bigvee_{n\geq m} a_n \quad \text{(maximum limit)}$$

$$\liminf_{n\to\infty} a_n = \bigvee_{m\in\mathbb{N}} \bigwedge_{n\geq m} a_n \quad \text{(minimum limit)}$$

- Let $(E_n)_{n\in\mathbb{N}}$ be a sequence of events in the measurable space $(\Omega, \mathcal{F})$.

$$\limsup_{n\to\infty} E_n = \bigcap_{m\in\mathbb{N}} \bigcup_{n\geq m} E_n \quad (E_n \text{ infinitely often})$$

$$\liminf_{n\to\infty} E_n = \bigcup_{m\in\mathbb{N}} \bigcap_{n\geq m} E_n \quad (E_n \text{ eventually})$$

A similar definition applies to sequences of functions. If $(f_n)_n$ is a sequence of non-negative functions, then the set of $x \in S$ such that $\lim_n f_n(x) = 0$ is equal to the set $\{\limsup_n f_n = 0\}$.

Fatou lemma

Theorem

$$\mathrm{P}\left(\liminf_{n\to\infty} E_n\right) \leq \liminf_{n\to\infty} \mathrm{P}\left(E_n\right) \leq \limsup_{n\to\infty} \mathrm{P}\left(E_n\right) \leq \mathrm{P}\left(\limsup_{n\to\infty} E_n\right)$$

- $(\limsup_n E_n)^c = \liminf_n E_n^c$; $\limsup_n \mathbf{1}_{E_n} = \mathbf{1}_{\limsup_n E_n}$.
- Proof of FL. Write $\cup_m \cap_{n\geq m} E_n = \cup_m G_m$ so that $G_m \uparrow G = \liminf_n E_n$. We have $\mathrm{P}\left(G_m\right) \leq \wedge_{n\geq m} \mathrm{P}\left(E_n\right)$; monotone continuity (increasing) implies $\mathrm{P}\left(G_m\right) \uparrow \mathrm{P}\left(G\right)$ hence, $\vee_m \mathrm{P}\left(G_m\right) = \mathrm{P}\left(G\right)$. The middle inequality is a property of lim inf and lim sup. The least inequality follows from a similar proof using continuity on decreasing sequences or, by taking the complements.
- BC1. Assume $\sum_{n=1}^{\infty} \mathrm{P}\left(E_n\right) < +\infty$. We have for all $m \in \mathbb{N}$ that

$$\mathrm{P}\left(\limsup_n E_n\right) \leq \mathrm{P}\left(\cup_{n\geq m} E_n\right) \leq \sum_{n=m}^{\infty} \mathrm{P}\left(E_n\right) \to 0 \quad \text{if } m \to \infty$$

hence $\mathrm{P}\left(\limsup_n E_n\right) = 0$.

Measurable function

Definition

Given measurable spaces $(S_i, \mathcal{S}_i)$, $i = 1, 2$, we say that the function $h\colon S_1 \to S_2$ is measurable, or is a random variable, if for all $B \in \mathcal{S}_2$ the set $h^{-1}(B) = \{s \in S_1 | h(s) \in B\}$ belongs into $\mathcal{S}_1$.

Theorem

- *Let $\mathcal{C} \subset \mathcal{S}_2$ and $\sigma(C) = \mathcal{S}_2$. If $h^{-1}\colon \mathcal{C} \to \mathcal{S}_1$, then h is measurable.*
- *Given measurable spaces $(S_i, \mathcal{S}_i)$, $i = 1, 2, 3$, if both $h\colon S_1 \to S_2$, $g\colon S_2 \to S_3$ are measurable functions, then $g \circ f : S_1 \to S_3$ is a measurable function.*
- *Given measurable spaces $(S_i, \mathcal{S}_i)$, $i = 0, 1, 2$ and $h_i\colon S_0 \to S_j$, $j = 1, 2$, consider $h = (h_1, h_2)\colon S_0 \to S_1 \times S_2$. with product space $(S_1 \times S_2, \mathcal{S}_1 \otimes \mathcal{S}_2)$, Then both h_1 and h_2 are measurable if, and only if, h is measurable.*

- Ch. 3 of Ch. 2 of .

Image measure

Definition

Given measurable spaces $(S_i, \mathcal{S}_i)$, $i = 1, 2$, a measurable function $h\colon S_1 \to S_2$, and a measure μ_1 on $(S_1, \mathcal{S}_1)$, then $\mu_2 = \mu_1 \circ h^{-1}$ is a measure on $(S_2, \mathcal{S}_2)$. We write $h_\# \mu_1 = \mu_2 \circ h^{-1}$ and call it image measure. If μ_1 is a probability measure, we say that $h_\# \mu_1$ is the distribution of the random variable h.

- Bernoulli scheme Let $(\Omega, \mathcal{F}, \mathrm{P})$ be the Bernoulli scheme, and define $X_t\colon \Omega \to \{0, 1\}$ to be the t-projection, $X_t(x_1 x_2 \cdots) = x_t$. It is a random variable with Bernoulli distribution $\mathrm{B}(\theta)$. The random variable $Y_n = X_1 + \cdots + X_n$ has distribution $\mathrm{Bin}(\theta, n)$. The random variable $T = \inf\{k \in \mathbb{N} | X_k = 1\}$ has distribution $\mathrm{Geo}(\theta)$.

Real random variable

Definition

Let $(S, \mathcal{S})$ be a measurable space. A real random variable is a real function $h\colon S \to \mathbb{R}$ with is measurable into $(\mathbb{R}, \mathcal{B})$.

Theorem

- *$h\colon S \to \mathbb{R}$ is a real random variable if, and only if, for all $c \in \mathbb{R}$ the level set $\{s \in S\}\, h(s) \leq c$ is measurable. The same property holds with $\leq$ replaced by $<$ or $\geq$ or $>$. The condition can be taken as a defintion of extended random variable i.e.*
  $h\colon S \to \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$.
- *If $g, h\colon S \to \mathbb{R}$ are real random variables and $\Phi\colon \mathbb{R}^2 \to \mathbb{R}$ is continuous, then $\Phi \circ (g, h)$ is a real random variable.*
- *Let $(h_n)_{n\in\mathbb{N}}$ be a sequence of real random variables on $(S, \mathcal{S})$. Then $\sup_n f_n$, $\inf_n f_n$, $\limsup_n f_n$, $\liminf_n f_n$ are real random variable.*

A monotone-class theorems

Theorem

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

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

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

- Application. Consider measurable spaces $(\Omega_i, \mathcal{F}_i)$, $i = 1, 2$. Define $\Omega = \Omega_1 \times \Omega_2$ and $\mathcal{I} = \{A_1 \times A_2 | A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2\}$. Then $\mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma(\mathcal{I})$. Let $\mathcal{H}$ be the set of all bounded real funtions $f : \Omega_1 \times \Omega_2 \to \mathbb{R}$ such that for each fixed $x \in \Omega_1$ the mapping $\Omega_2 \ni y \mapsto f(x, y)$ is $\mathcal{F}_2$-measurable and for each fixed $y \in \Omega_2$ the mapping $\Omega_1 \ni x \mapsto f(x, y)$ is $\mathcal{F}_1$-measurable.

- §3.14 and §A3.1 of ; Hendout 1.