MATH60028 Probability Theory

Lectured by Professor Igor Krasovsky

Autumn 2023
Notes transcribed into LaTeX by James Chen

1 Events, probability, random variables

Let Ω be a set of points ω.

Definition 1.1. A nonempty system of subsets of Ω is called an algebra $\mathcal{A}$ if $\Omega \in \mathcal{A}$, $A \cup B$, $A \cap B$, $A^c = \Omega \setminus A$ are elements of $\mathcal{A}$ whenever $A, B \in \mathcal{A}$.

Definition 1.2. A function $\mu : \mathcal{A} \to [0, \infty]$ is called finitely additive measure if for any $\underline{\text{disjoint}}$ $A, B \in \mathcal{A}$:

- $\mu(A \cup B) = \mu(A) + \mu(B)$

Note that then $\forall A, B \in \mathcal{A}$, we have: $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$.

Definition 1.3. An algebra $\mathcal{A}$ is called a σ-algebra $\mathcal{F}$ if a countable union $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$ whenever $A_1, A_2, \cdots \in \mathcal{F}$.

Note that then also $\bigcap_{n=1}^{\infty} A_n \in \mathcal{F}$: consider $\Omega \setminus \hat{A}_k$.

Definition 1.4. A function $\mu : \mathcal{F} \to [0, \infty]$ is called σ-additive if for any $\underline{\text{disjoint}}$ $A_1, A_2, \cdots \in \mathcal{F}$, we have $\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$. Such μ is called a **measure** on $\mathcal{F}$. A measure μ is ccalled a **probability measure** if $\mu(\Omega) = 1$.

Note that $\mu(\emptyset) = 0$ since $\mu(\emptyset) = \mu(\emptyset \cup \emptyset) = 2\mu(\emptyset)$.

A measure is called σ-finite if there exists a representation $\Omega = \bigcup_{k=1}^{\infty} \Omega_k$, Ω_k are pairwise disjoint, $\mu(\Omega_k) \leq \infty$ for all $k = 1, 2, \cdots$.

Definition 1.5. A **probability space** is a triple $(\Omega, \mathcal{F}, \mathbb{P})$ where Ω is a set called the sample space, $\mathcal{F}$ is a σ-algebra of subsets of Ω, $\mathbb{P}$ a probability measure on $\mathcal{F}$. Any element of $\mathcal{F}$ is called an event.

We also say:

- $\mathbb{P}(A \cup B)$ - probability that either A or B occurs.
- $\mathbb{P}(A \cap B)$ - probability that both A and B occur.
- $\mathbb{P}(A^c) = \mathbb{P}(\Omega \setminus A)$ - probability that A does not occur.

Lemma 1.1 (Continuity of measure).
(a) If $A_n \in \mathcal{F}$, $A_1 \subset A_2 \subset \cdots$, then:

$$\mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = \lim_{n \to \infty} \mathbb{P}(A_n) \tag{1.1}$$

- continuity from below

(b) If $B_n \in \mathcal{F}$, $B_1 \supset B_2 \supset \cdots$, then:

$$\mathbb{P}(\bigcap_{n=1}^{\infty} B_n) = \lim_{n \to \infty} \mathbb{P}(B_n) \tag{1.2}$$

- continuity from above

Proof. Since $\bigcup_n A_n = (\bigcap_n A_n^c)^c$, (a) and (b) are equivalent. So sufficient to show (b). To show (b), it suffices to show it for the case $\bigcap_{n=1}^{\infty} B_n = \emptyset$, as otherwise we replace B_n by $B_n \setminus \bigcap_{n=1}^{\infty} B_n = \emptyset$.

Remark 1.1.1. (b) with $\bigcap_{n=1}^{\infty} B_n = \emptyset$ is called **continuity at zero.**

To prove continuity at zero *i.e.*the fact that $\lim_{n\to\infty} \mathbb{P}(B_n) = 0$, consider sets $\Omega \setminus B_1, B_1 \setminus B_2, \cdots$. They are disjoint.

$$\emptyset = \bigcap_{n=1}^{\infty} B_n = \Omega \setminus \bigcup_{n=1}^{\infty} (\Omega \setminus B_n) \tag{1.3}$$

$$= \Omega \setminus ((\Omega \setminus B_1) \cup (B_1 \setminus B_2) \cup \cdots) \tag{1.4}$$

Therefore,

$$1 = \mathbb{P}(\Omega \setminus B1) + \mathbb{P}(B_1 \setminus B2) + \cdots \tag{1.5}$$

$$= \sum_{j=0}^{\infty} \mathbb{P}(B_j \setminus B_{j+1}) \tag{1.6}$$

where we have set $B_0 = \Omega$. So $\forall \varepsilon > 0, \exists n_0$ s.t. $\forall n > n_0$, we have:

$$1 - \sum_{j=0}^{n-1} \mathbb{P}(B_j \setminus B_{j+1}) = \mathbb{P}(B_n) \leq \varepsilon \tag{1.7}$$

□

Any finitely additive measure μ on $\mathcal{F}$ satisfies the property:

$$\sum_{n=1}^{\infty} \mu(A_n) \leq \mu\left(\bigcup_{n=1}^{\infty} A_n\right) \tag{1.8}$$

if A_j are disjoint sets in $\mathcal{F}$.
When is it also a measure (equality holds)?

Lemma 1.2 (finite additivity into countable additivity)**.**
A finitely additive probability measure on $\mathcal{F}$ is a probability measure iff it is continuous at zero.

Proof. ($\Rightarrow$) if $\mathbb{P}$ is a probability measure then continuity at zero is already shown.

($\Leftarrow$) Let $\mathbb{P}$ be finitely additive probability measure on $\mathcal{F}$ and for any $B_1 \supset B_2 \supset \cdots$, $\bigcup_{n=1}^{\infty} B_n = \emptyset$, $B_n \in \mathcal{F}$, we have $\lim_{n\to\infty} \mathbb{P}(B_n) = 0$.
Hence (a) of Lemma 1.1 holds since (a) and (b) are both equivalent to continuity at zero.
For any disjoint $C_1, C_2, \cdots \in \mathcal{F}$, define $A_n = \bigcup_{k=1}^{n} C_k$. Then $A_1 \subset A_2 \subset \cdots$ and by (a):

$$\mathbb{P}\left(\bigcup_{k=1}^{\infty} C_k\right) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mathbb{P}(A_n) \tag{1.9}$$

$$= \lim_{n\to\infty} \sum_{k=1}^{n} \mathbb{P}(C_k) = \sum_{k=1}^{\infty} \mathbb{P}(C_k) \tag{1.10}$$

Here (a) was used in the second equality and finite additivity was used on the third equality. □

Examples of σ-algebras on Ω:

- $\mathcal{F}_* = \{\emptyset, \Omega\}$
- $\mathcal{F}^* = \{A : A \subset \Omega\} = 2^{\Omega}$ (Power set)

- σ-algebra generated by partitions: $\sigma(D) = \left\{\bigcup_{j\in I} D_j : I \subset \mathbb{N}\right\}$ where $D = \{D_1, D_2, \cdots\}$ is a partition of Ω into a countable union of disjoint D_j, $\Omega = \bigcup_{j=1}^{\infty} D_j$.

Lemma 1.3. *For any collection $\mathcal{E}$ of subsets of Ω, there exists a minimal algebra $a(\mathcal{E})$ and a minimal σ-algebra $\sigma(\mathcal{E})$ that contains all elements of $\mathcal{E}$ (intersection of all algebras (resp. σ-algebras containing $\mathcal{E}$)).*

Proof. Recall that the intersection of arbitrarily many algebras (resp. σ-algebras) containing $\mathcal{E}$ is an algebra (resp. σ-algebras) containing $\mathcal{E}$. We say that $\sigma(\mathcal{E})$ is generated by $\mathcal{E}$. □

Definition 1.6. A measurable space is a pair $(E, \mathcal{E})$ where E is a set and $\mathcal{E}$ is a σ-algebra on E.

Examples: (1) $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, where $\mathcal{B}(\mathbb{R}^n) = \sigma(\{A \subset \mathbb{R}^n : A \text{open}\})$.
For $n = 1$, recall that $\mathcal{B} = \sigma(\{\text{open subsets of } \mathbb{R}\})$

$$= \sigma(\{\text{open intervals}\}) = \sigma(\{\text{closed intervals}\}) = \sigma(\{\text{closed half-lines } (-\infty, x], x \in \mathbb{R})\}) \tag{1.11}$$

(2) Product σ-algebra
Consider $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$ where $(\Omega_j, \mathcal{F}_j)$ are measurable spaces, $j = 1, 2$.

$$\mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma(\{A_1 \times A_2 : A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2\}) \tag{1.12}$$

Lemma 1.4. $\mathcal{B}(\mathbb{R}^2) = \mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R})$.

Proof. $\mathcal{B}(\mathbb{R}^2) \subset \mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R})$, since any open $A \subset \mathbb{R}^2$ can be written as follows:

$$A = \bigcup_{x \in A \cap \mathbb{Q}^2} R(x, \tau(x)) \in \mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R}) \tag{1.13}$$

where $R(x, \tau)$ is an open square centered at x of side length τ.
Other direction: sufficient to check that $B_1 \times B_2 \in \mathcal{B}(\mathbb{R}^2)$ for any 2 borel sets B_1 and B_2.
Note that $B_1 \times \mathbb{R} \in \mathcal{B}(\mathbb{R}^2)$ since

$$B_1 \times \mathbb{R} = \sigma(\{\text{open subsets of } \mathbb{R}\}) \times \mathbb{R} = \sigma(\{\text{open subsets of } \mathbb{R}\} \times \mathbb{R}) \tag{1.14}$$

Similarly, $\mathbb{R} \times B_2 \in \mathcal{B}(\mathbb{R}^2)$, and so $B_1 \times B_2 = (B_1 \times \mathbb{R}) \cap \mathbb{R} \times B_2 \in \mathcal{B}(\mathbb{R}^2)$ □

(3) Cylindrical σ-algebra
Let $\mathbb{R}^\infty = \{x = (x_1, x_2, \cdots), x_k \in \mathbb{R}\}$

Definition 1.7. A set $C \subset \mathbb{R}^\infty$ is called cylindrical if it is of the form $C = \{x \in \mathbb{R}^\infty : (x_1, x_2, \cdots, x_n) \in \tilde{C}_n\}$ for some $n \geq 1$ and $\tilde{C}_n \in \mathcal{B}(\mathbb{R}^n)$.

Cylindrical sets form an algebra (check!) which generates a σ-algebra called cylindrical σ-algebra, denoted $\mathcal{B}(\mathbb{R}^\infty)$.
One can verify that $\mathcal{B}(\mathbb{R}^n) = \sigma(\{A_1 \times A_2 \times \cdots \subset \mathbb{R}^\infty, A_k \in \mathcal{B}(\mathbb{R})\})$
Example:
$\forall c \in \mathbb{R}$, let:

$$A = \{x \in \mathbb{R}^\infty : \limsup_{n\to\infty} x_n > c\} \tag{1.15}$$

We have that $A \in \mathcal{B}(\mathbb{R}^\infty)$: indeed

$$A = \bigcap_{n=1}^{\infty} \bigcup_{k=n+1}^{\infty} \{x \in \mathbb{R}^\infty : x_k > c\} = \{x_k > c \text{ i.o.}\} \tag{1.16}$$

$\forall c, D = \{x \in \mathbb{R}^\infty : \lim_{n\to\infty} x_n = c\} \in \mathcal{B}(\mathbb{R}^\infty)$ (Exercise!)
Recall: nondecreasing function $g(x)$ on $\mathbb{R}$ is continuous up to possibly countably many discontinuities of first kind: $f(x+0) and f(x-0)$ both exist but $f(x+0) - f(x-0) = h_x > 0$.
Moreover, the derivative $g'(x)$ exists Lebesgue a.e.
Let $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P})$ - probability space, $F(x) = \mathbb{P}((-\infty, x]), x \in \mathbb{R}$. Then (exercise):

- $F(x)$ is non decreasing
- $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to\infty} F(x) = 1$
- $F(x)$ is continuous from the right $\forall x \in \mathbb{R}$

Definition 1.8. Any function $F : \mathbb{R} \to [0,1]$ satisfying the above 3 conditions is called a distribution function (on $\mathbb{R}$).

We have seen that $\mathbb{P}$ gives rise to F. In fact, the opposite is true and there exists a one-to-one correspondence between distribution functions and probability measures:

Theorem 1.5. *Let $F(x)$ be a distribution function on $\mathbb{R}$.*
Then there exists a unique probability measure $\mathbb{P}$ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ s.t. $\mathbb{P}((-\infty, x])) = F(x)$.