STAT600 - Probability theory I

Haoran Li

2020/08/18

Contents

1 σ-algebras

Definition 1.1. Ω is a set. $\mathcal{F}$ is called an *algebra* if

- $\varnothing, \Omega \in \mathcal{F}$
- $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
- $A_1, A_2 \in \mathcal{F} \Rightarrow A_1 \cup A_2 \in \mathcal{F}$

An algebra $\mathcal{F}$ is called a *σ-algebra* if it satisfies σ additivity

- $\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{F} \Rightarrow \displaystyle\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

Note. If $\mathcal{F}$ is a σ-algebra, $A \setminus B = (A^c \cup B)^c \in \mathcal{F}$, $\displaystyle\bigcap_{i=1}^{\infty} A_i = \left(\bigcup_{i=1}^{\infty} A_i^c\right)^c \in \mathcal{F}$

Note. $(\Omega, \mathcal{F})$ is called a *measurable space*

Proposition 1.2. Every finite algebra $\mathcal{F}$ has 2^n elements for some n

Example 1.3. Consider $R_{a,b,c,d} = (a, b] \times (c, d]$, $0 \leq a \leq b \leq 1, 0 \leq c \leq d \leq 1$, $\mathcal{F}$ is the set of all finite unions of $R_{a,b,c,d}$'s, $\Omega = (0, 1] \times (0, 1]$. $\mathcal{F}$ is an algebra but not a σ-algebra, just consider $R_{a,b,c,d}^c$

Definition 1.4. $(\Omega, \mathcal{F})$ is a measurable space, $P : \mathcal{F} \to \mathbb{R}$ is a *probability measure* if

- $P(A) \geq 0$ for all $A \in \mathcal{F}$
- $P(\Omega) = 1$
- $P\left(\displaystyle\bigsqcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$

Elements in $\mathcal{F}$ are called *events*

Remark 1.5. If we only assume $P(A_1 \sqcup A_2) = P(A_1) + P(A_2)$, then we still have $P\left(\displaystyle\bigsqcup_{i=1}^{\infty} A_i\right) \geq \sum_{i=1}^{\infty} P(A_i)$

Lemma 1.6.

1. If $A_1 \subseteq A_2 \subseteq \cdots$, then $P(\bigcup A_i) = \lim_{n\to\infty} P(A_n)$
2. If $A_1 \subseteq A_2 \subseteq \cdots$, then $P(\bigcup A_i) = \lim_{n\to\infty} P(A_n)$
3. σ additivity

1. $\Leftrightarrow$ 2. $\Leftrightarrow$ 3.

Proposition 1.7 (Inclusion-exclusion inequality)**.**

$$P(A_1 \cup \cdots \cup A_n) \leq \sum_{i=1}^{n} P(A_i)$$

$$P(A_1 \cup \cdots \cup A_n) \geq \sum_{i=1}^{n} P(A_i) - \sum_{i_1 < i_2} P(A_{i_1} \cap A_{i_2})$$

And so on

Definition 1.8. $\mathcal{G} \subseteq \mathscr{P}(\Omega)$, $\sigma(\mathcal{G})$ is the minimal σ-algebra that contains all elements of $\mathcal{G}$

Definition 1.9. The *Borel algebra* $\mathscr{B}(X)$ is the minimal σ-algebra generated by all open subsets of X

Definition 1.10. $(\Omega_1, \mathcal{F}_1), (\Omega_2, \mathcal{F}_2)$ are measurable spaces, the *product space* is $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2)$ where $\mathcal{F}_1 \times \mathcal{F}_2$ is really a short hand for $\sigma(\mathcal{F}_1 \times \mathcal{F}_2)$

Definition 1.11. A topological space X is called *separable* if it contains a countable dense subset

Theorem 1.12. $(M_1, d_1), (M_2, d_2)$ are separable metric spaces. $d(x, y) = \sqrt{d_1(x, y)^2 + d_2(x, y)^2}$, then $\mathscr{B}(M) = \mathscr{B}(M_1) \times \mathscr{B}(M_2)$

Remark 1.13. $\mathscr{B}(X)$ is generally bigger than the minimal σ-algebra generated by open balls, a counter example would be a discrete metric space, however this is true if X is a separable metric space

Definition 1.14. $f : \Omega \to \Omega'$ is *measurable* if $f^{-1}(A) \in \mathcal{F}$ for all $A \in \mathcal{F}'$. A *random variable* is a measurable function $f : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathscr{B}(\mathbb{R}))$, or equivalently $f^{-1}(-\infty, a] \in \mathcal{F}$

Definition 1.15. $F : \mathbb{R} \to [0, 1]$ is a *distribution function*

- F is non-decreasing
- $\lim\limits_{x \to +\infty} F(x) = 1$, $\lim\limits_{x \to -\infty} F(x) = 0$
- For any $x \in \mathbb{R}$, $\lim\limits_{y \searrow x} F(y) = F(x)$

$F_\xi(x) = P(\{\omega | \xi(\omega) \leq x\}) = P(\xi \leq x)$ is a distribution function, conversely, given a distribution function F, there exists a random variable ξ such that $F = F_\xi$

Definition 1.16. The *mathematical expectation* $E\xi = \int_\Omega \xi dP$ provided $\int_\Omega |\xi| dP < \infty$

Theorem 1.17 (Chebychev's inequality)**.**

$$P(|\xi - E\xi| \geq c) = P(|\xi - E\xi|^2 \geq c^2) \leq \frac{\operatorname{Var} \xi}{c^2}$$

Definition 1.18. $f : (\Omega, \mathcal{F}, P) \to (\Omega', \mathcal{F}')$ is measurable, the *induced measure* P' is such that $P'(A) = P(f^{-1}(A))$. If $\xi : \Omega' \to \mathcal{R}$ is a random variable, then $\int_{\Omega'} \xi dP' = \int_\Omega \xi \circ f dP$ is change of variable

Exercise 1.19. Chapter 1: 5,6,14
Chapter 3: 2,3,4,5,6,7

Exercise 1.20. ξ_n are random variables, F is the distribution, show that

- $A = \left\{ \omega \,\middle|\, \lim\limits_{n \to \infty} \xi_n(\omega) \text{ exists} \right\} \in \mathcal{F}$
- $\displaystyle\int_{-\infty}^{\infty} F(x + 10) - F(x) dx = 10$

Theorem 1.21. Let $\mathcal{I} = \{(a, b], (-\infty, b], (a, \infty), (-\infty, \infty)\}$, suppose $m : \mathcal{I} \to \mathbb{R}$ is a function such that $m(I) \geq 0$ for all $I \in \mathcal{I}$, $m\left(\bigsqcup\limits_{i=1}^{\infty} I_i\right) = \sum\limits_{i=1}^{\infty} m(I_i)$ for $\{I_i\} \subseteq \mathcal{I}$, then there exists a unique measure μ on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ such that $\mu(I) = m(I)$ for all $I \in \mathcal{I}$

Proof. Suppose F is a distribution, define $m((a, b]) = F(b) - F(a)$, then there exists a unique measure μ on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ such that $\mu((-\infty, b]) = F(b)$, $\xi : (\mathbb{R}, \mathscr{B}(\mathbb{R}), \mu) \to \mathbb{R}$, $\xi(x) = x$, then $\mu(\xi \leq b) = F(b)$ ∎

Definition 1.22. F is a distribution function, g is measurable on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, define $\int_{\mathbb{R}} gdF = \int_{\mathbb{R}} gd\mu_F$

Definition 1.23. A measure μ is *σ-finite* if $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) < \infty$ for $\{A_n\} \subseteq \mathcal{F}$

Definition 1.24. A measure μ is *locally finite* if there exist $\Omega_1, \Omega_2, \cdots$ such that $\Omega_n \subseteq \Omega_{n+1}$, $\bigcup \Omega_i = \Omega$, $\mu(\Omega_n) < \infty$

Definition 1.25. A measure μ on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ is *discrete* if $A = \{a_1, a_2, \cdots\}$, finite or countable, such that $\mu(\mathbb{R}) = \mu(A)$

Definition 1.26. A measure μ on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ is *singular continuous* if there exists $A \subseteq \mathbb{R}$ such that $m(A) = 0$, $\mu(\mathbb{R}) = \mu(A)$ and $\mu(\{r\}) = 0$ for all $r \in R$

Theorem 1.27. If μ is a measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, then there exist unique measures μ_1, μ_2, μ_3 such that $\mu = \mu_1 + \mu_2 + \mu_3$, μ_1 is discrete, μ_2 is singular continuous and μ_3 is absolutely continuous

Proof. Let $a_1^1, \cdots, a_{k_1}^1$ be all the points such that $\mu(a_i^1) \geq 1$, $a_1^2, \cdots, a_{k_2}^2$ be all the points such that $\mu(a_i^2) \geq \frac{1}{2}$, and so on. Let $A = \{a_i^j\}$, define $\mu_1(B) = \mu(A \cap B)$, $\mu' = \mu - \mu_1$, for any a, $\mu'(\{a\}) = 0$. Find a Borel set A_1 such that $\lambda(A_1) = 0$, $\mu'(A_1) \geq \frac{1}{k}$ with smallest possible k, if no such k exists, take $A_1 = \varnothing$. Find a Borel set $A_2 \subseteq (\mathbb{R} \setminus A_1)$ such that $\lambda(A_2) = 0$, $\mu'(A_2) \geq \frac{1}{k}$ with smallest possible k, and so on. $A' = \bigcup A_i$. Uniqueness ☐

Definition 1.28. ρ is the *density* of a distribution function F if $F(b) - F(a) = \int_a^b \rho(t)d\lambda(t)$, by the uniqueness of the extension theorem, $\mu_F(A) = \int_A \rho d\lambda$. p is the density of ξ if $P(\xi \in A) = \int_A pd\lambda$

Example 1.29. C is the Cantor set, F is Cantor function, with $F(x) = \lim_{\substack{y \searrow x \\ y \in C}} F(y)$, then F is continuous

References

[1] *Theory of probability and random processes (second edition)* - Leonid B. Koralov

Index