Math 6261 23-01-13

1. Measure theory

1.1 Measure spaces and probability spaces

Def. Let $\Omega \neq \emptyset$. A collection $\mathcal{F}$ of subsets of Ω is said to be a σ-algebra on Ω if

(1) If $F \in \mathcal{F}$, then $F^c \in \mathcal{F}$, where $F^c = \Omega \setminus F$

(2) If $F_n \in \mathcal{F}$, $n \geq 1$, then $\bigcup_{n=1}^{\infty} F_n \in \mathcal{F}$.

Remark: • A σ-algebra is closed under the complement, countable union and countable intersection

$$\left(\bigcap_{n=1}^{\infty} F_n = \left(\bigcup_{n=1}^{\infty} F_n^c \right)^c \right)$$

• Let $\{\mathcal{F}_i\}_{i \in \mathcal{I}}$ be a family of σ-algebras on Ω.
Then

$$\bigcap_{i \in \mathcal{I}} \mathcal{F}_i \quad \text{is also a } \sigma\text{-algebra on } \Omega.$$

Def. Let $\mathcal{A}$ be a collection of subsets of Ω.
Let $\sigma(\mathcal{A})$ denote the smallest σ-algebra on Ω that contains $\mathcal{A}$.
We call $\sigma(\mathcal{A})$ the σ-algebra generated by $\mathcal{A}$.

Example • Let X be a topological space. Let $\mathcal{B}(X)$ denote the σ-algebra generated by the collection of open sets in X.
We call $\mathcal{B}(X)$ the Borel σ-algebra.

Let $\mathcal{B}(\mathbb{R})$ denote the Borel σ-algebra on $\mathbb{R}$. Each element in $\mathcal{B}(\mathbb{R})$ is called a Borel set in $\mathbb{R}$.

Def. (Measurable space) $(\Omega, \mathcal{F})$ is called a measurable space if $\Omega \neq \phi$ and $\mathcal{F}$ is a σ-algebra on Ω.

Def. (measure) A function $\mu : \mathcal{F} \to [0, \infty]$ is called a measure on $(\Omega, \mathcal{F})$ if

(i) $\mu(\phi) = 0$.

(ii) $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$ if

A_n, $n \geq 1$, are disjoint elements in $\mathcal{F}$.

Prop 1.1. Let μ be a measure on $(\Omega, \mathcal{F})$. Then

(i) $\mu(A) \leq \mu(B)$ if $A \subset B$ (monotonicity)

(ii) $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n)$ for $A_n \in \mathcal{F}$. (Sub-additivity)

(iii) If $A_n \uparrow A$, then $\mu(A_n) \to \mu(A)$ as $n \to \infty$ (continuity from below)

(iv) If $A_n \downarrow A$ and $\mu(A_1) < \infty$, then

$$\lim_{n \to \infty} \mu(A_n) = \mu(A).$$ (continuity from above).

Def. • A triple $(\Omega, \mathcal{F}, \mu)$ is called a measure space if μ is a measure on $(\Omega, \mathcal{F})$.

• If $\mu(\Omega) = 1$, we call μ a prob. measure. Correspondingly,

$(\Omega, \mathcal{F}, \mu)$ is called a prob. space.

Usually a prob. measure is denoted as P.

Example. (discrete prob. space)

Let Ω be a countable set. Let

$$\mathcal{F} = 2^{\Omega} := \{ A : A \subset \Omega \}.$$

Then $(\Omega, \mathcal{F})$ is a measurable space.

Let $\{p(\omega)\}_{\omega \in \Omega}$ be a prob. vector, i.e. $p(\omega) \geq 0$ and $\sum_{\omega \in \Omega} p(\omega) = 1$.

Define

$$P(A) = \sum_{\omega \in A} p(\omega) \qquad \text{for all } A \subset \Omega.$$

Then $(\Omega, \mathcal{F}, P)$ is a (discrete) prob. space.

Example (Borel measure on $\mathbb{R}$) A measure μ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is called a Borel measure on $\mathbb{R}$.

Prop 1.2. Let μ be a Borel prob. measure on $\mathbb{R}$. Set

$$F(x) = \mu\big((-\infty, x]\big) \quad \text{for} \quad x \in \mathbb{R}.$$

Then

(1) F is non-decreasing, i.e. $F(x) \leq F(y)$ if $x < y$.

(2) F is right-continuous, i.e.

$$\lim_{y \to x+} F(y) = F(x).$$

(3) $\lim_{x \to +\infty} F(x) = 1$, $\quad \lim_{x \to -\infty} F(x) = 0$.

Pf. (1) is trivial. (2) & (3) follow from the continuity property of a prob. measure. $\blacksquare$

1.2 Random variables and their distributions.

Let $(\Omega, \mathcal{F}, P)$ be a probability space.

Def. A function $X : \Omega \to \mathbb{R}$ is said to be $\mathcal{F}$-measurable if

$$X^{-1}(A) \in \mathcal{F} \quad \text{for every Borel set } A \subset \mathbb{R}.$$

If so, we call X a random variable (r.v.).

Example: • Let $(\Omega, \mathcal{F}, P)$ be a discrete prob. space.
Then any function $X : \Omega \to \mathbb{R}$ is a r.v.

- Let $(\Omega, \mathcal{F}, P)$ be a general prob. space and let $A \in \mathcal{F}$.

Define $\mathbb{1}_A : \Omega \to \mathbb{R}$ by

$$\mathbb{1}_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\mathbb{1}_A$ is a r.v. which is called the indicator function of A.

(check: $\mathbb{1}_A^{-1}\{1\} = A$, $\mathbb{1}_A^{-1}\{0\} = A^c$)

Def. Let X be a r.v. on $(\Omega, \mathcal{F}, P)$. Then X induces a prob. measure μ on $\mathbb{R}$ by

$$\mu(A) = P(X \in A) := P(X^{-1}(A)), \quad A \in \mathcal{B}(\mathbb{R}).$$

We call μ the distribution of X.

Moreover, set $F(x) = P\{X \leq x\} = \mu((-\infty, x])$ for $x \in \mathbb{R}$; we call it the distribution function of X.

If $F(x)$ has the form

$$F(x) = \int_{-\infty}^{x} f(y)\, dy,$$

then we say X has the density function f.

Example: ① (Uniform distribution on $(0,1)$)

- $f(x) = \begin{cases} 1 & \text{for } x \in (0,1) \\ 0 & \text{otherwise.} \end{cases}$

- $F(x) = \begin{cases} 1 & \text{if } x \geq 1 \\ x & \text{if } 0 \leq x \leq 1 \\ 0 & \text{if } x < 0. \end{cases}$

② (exponential distribution with parameter λ)

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{if } x > 0 \\ 0 & \text{if } x \leq 0. \end{cases}$$

③ (standard norm distribution)

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}, \quad x \in \mathbb{R}.$$

1.3. Random elements and random vectors.

Now we generalize the concept of r.v.

Def. A map $X : \Omega \to T$ from $(\Omega, \mathcal{F}, P)$ to a measurable space $(T, \mathcal{T})$ is said to be measurable if

$$X^{-1}(A) \in \mathcal{F} \quad \text{for every } A \in \mathcal{T}.$$

In this case, we call X a random element of $(T, \mathcal{T})$.

If $(T, \mathcal{T}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, then we call X a random vector.

Def. Let $X : \Omega \to T$ be a random element. Set

$$\sigma(X) = \{ X^{-1}(A) : A \in \mathcal{T} \}.$$

We call it the σ-algebra generated by X.

Below we give a useful result to check the measurability of $X : \Omega \to T$.

Prop 1.3. Let $X : \Omega \to T$ be a map. Suppose $\mathcal{A}$ is a collection of subsets of T such that $\sigma(\mathcal{A}) = \mathcal{T}$. Then

$$X \text{ is } \mathcal{F}\text{-measurable} \iff X^{-1}(A) \in \mathcal{F} \text{ for all } A \in \mathcal{A}.$$

Pf. "$\Leftarrow$" Let $\mathcal{D} := \{ A \subset T : X^{-1}(A) \in \mathcal{F} \}$. Then $\mathcal{D}$ is a σ-algebra and contains $\mathcal{A}$. Hence $\mathcal{D} \supset \sigma(\mathcal{A}) = \mathcal{T}$. $\blacksquare$.

Prop 1.4. If $X : (\Omega, \mathcal{F}) \to (T, \mathcal{T})$ and $f : (T, \mathcal{T}) \to (U, \mathcal{U})$ are measurable, then so is $f(X) : (\Omega, \mathcal{F}) \to (U, \mathcal{U})$.

Pf. Let $A \in \mathcal{U}$. Then $f^{-1}(A) \in \mathcal{T}$. Thus

$$X^{-1}\left(f^{-1}(A)\right) \in \mathcal{F}.$$

Hence $\left(f(X)\right)^{-1}(A) = X^{-1}\left(f^{-1}(A)\right) \in \mathcal{F}$. $\blacksquare$

- Extended real line $\mathbb{R}^* = [-\infty, \infty]$.

  Endow $\mathbb{R}^*$ with the topology generated by

$$[-\infty, a), \quad (a, b), \quad (b, +\infty]$$

Let $\mathcal{B}(\mathbb{R}^*)$ denote the Borel σ-algebra on $\mathbb{R}^*$.

A measurable map $X : (\Omega, \mathcal{F}) \to (\mathbb{R}^*, \mathcal{B}(\mathbb{R}^*))$ is also called a random variable.

Prop. 1.5. Let $X_1, X_2, \cdots,$ be r.v.'s. Then

$$\inf_n X_n, \quad \sup_n X_n, \quad \underline{\lim}_n X_n, \quad \overline{\lim}_n X_n$$

are all r.v.'s.

1.4 Integration

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f : \Omega \to \mathbb{R}^*$ be measurable.

Then we can define

$$\int f \, d\mu = \int f^+ d\mu - \int f^- d\mu$$

if one of $\int f^+ d\mu$, $\int f^- d\mu$ is finite.

We call f integrable if $\int |f| \, d\mu < \infty$, and write

$f \in L^1(\Omega, \mathcal{F}, \mu)$ or $L^1(\mu)$.

Moreover we write $f \in L^p(\mu)$ if $\int |f|^p d\mu < \infty$

and $\|f\|_p := \left(\int |f|^p d\mu \right)^{1/p} \longrightarrow$ p norm of f

Basic inequalities:

Hölder inequality: Let $p, q >$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. Then

$$\int |fg| \, d\mu \le \left(\int |f|^p d\mu \right)^{1/p} \left(\int |g|^q d\mu \right)^{1/q}.$$

Minkowski inequality:

$$\|f + g\|_p \le \|f\|_p + \|g\|_p \quad \text{for all } p \ge 1.$$

Jensen inequality: Let μ be a prob. measure. Let $\varphi : \mathbb{R} \to \mathbb{R}$ be convex, i.e.

$$p\varphi(x) + (1-p)\varphi(y) \geq \varphi(px + (1-p)y)$$

for all $0 \leq p \leq 1$ and $x, y \in \mathbb{R}$. Suppose f and $\varphi(f)$ are integrable. Then

$$\varphi\left(\int f \, d\mu\right) \leq \int \varphi \circ f \, d\mu.$$

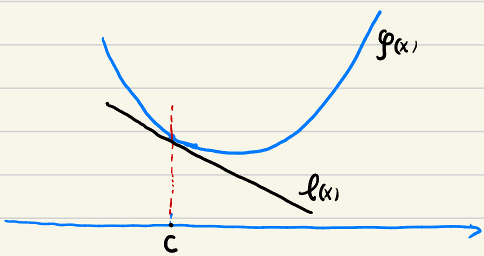
Pf. Write $c = \int f \, d\mu$.

Since φ is convex, there exists a function

$$\ell(x) = ax + b$$

such that $\ell(c) = \varphi(c)$ and $\varphi(x) \geq \ell(x)$ for all $x \in \mathbb{R}$.

See the following picture.



Hence

$$\varphi(f(x)) \geq \ell(f(x)) = af(x) + b$$

Taking integration gives

$$\int \varphi \circ f \, d\mu \geq \int (af(x) + b) \, d\mu = a \int f \, d\mu + b$$

$$= \ell(c) = \varphi(c). \quad \square$$

Next we recall some convergence results:

- (Monotone convergence Thm) Let f_n, $n \geq 1$, be non-negative functions such that $f_n \nearrow f$ a.e. Then

$$\int f_n \, d\mu \to \int f \, d\mu \quad \text{as } n \to \infty$$

- Fatou's lemma: Let f_n, $n \geq 1$, be non-negative measurable functions. Then

$$\varliminf_{n \to \infty} \int f_n \, d\mu \geq \int \varliminf_{n \to \infty} f_n \, d\mu.$$

- Dominated convergence Thm:

Suppose $f_n \to f$ a.e and $|f_n| \leq g$ for all n and $\int g \, d\mu < \infty$. Then

$$\int f_n \, d\mu \longrightarrow \int f \, d\mu.$$