Probability Notes 2024

Ruan Yuanlong

BUAA, Beijing, China
Email address: ruanyl@buaa.edu.cn

Contents

1. 单调类定理

Review:

- $\mathscr{A}$ is a field, $\mathscr{M}$ is a monotone class. Then

$$\mathscr{A} \subset \mathscr{M} \Longrightarrow \sigma(\mathscr{A}) \subset \mathscr{M}.$$

- $\mathscr{P}$ is a π-system, $\mathscr{L}$ is a λ-system. Then

$$\mathscr{P} \subset \mathscr{L} \Longrightarrow \sigma(\mathscr{P}) \subset \mathscr{L}.$$

- measurable spaces $(E, \mathscr{F}_E), (F, \mathscr{F}_F)$, $f : (E, \mathscr{F}_E) \mapsto (F, \mathscr{F}_F)$. f is $\mathscr{F}_E/\mathscr{F}_F$-measurable if

$$\sigma(f) \triangleq f^{-1}(\mathscr{F}_F) \subset \mathscr{F}_E.$$

  Call it $\mathscr{F}_E$-measurable if

$$(F, \mathscr{F}_F) = (\mathbb{R}, \mathscr{B}(\mathbb{R})).$$

- $f : (E, \mathscr{F}_E) \mapsto (F, \sigma(\mathscr{E}))$, f is $\mathscr{F}_E/\sigma(\mathscr{E})$-measurable if

$$f^{-1}(\mathscr{E}) \subset \mathscr{F}_E.$$

Thm 1 (π-λ theorem)**.** *$\mathscr{P}$ is a π-system, $\mathscr{L}$ is a λ-system. If $\mathscr{P} \subset \mathscr{L}$, then $\sigma(\mathscr{P}) \subset \mathscr{L}$.*

Def 1 (**Simple function**)**.** *$i = 1, ..., n$, $A_i \in \mathscr{F}$ (pairwise) disjoint, $c_i \in \mathbb{R}$. f is (measurable) simple if $f = \sum_{i=1}^{n} c_i 1_{A_i}$.*

Alt. $i = 1, ..., n$, $A_i \in \mathscr{F}$, $c_i \in \mathbb{R}$ non-zero distinct, f is simple if $f = \sum_{i=1}^{n} c_i 1_{A_i}$.

$\triangleright$ 1. *$a, b \in \mathbb{R}$, g simple, then $af + bg$ simple*

Thm 2 (**Simple approximation**)**.** (1) *$f \geqslant 0$ measurable. There exist simple $\{f_n\}$, $0 \leqslant f_n \uparrow f$, uniform if f is bounded.*

(2) *f measurable. There exist simple $\{f_n\}$, $f_n \to f$, uniform if f is bounded.*

PROOF. **1**. Let

$$f_n = \frac{[2^n f]}{2^n} \wedge n = \sum_{i=0}^{n2^n-1} \frac{i}{2^n} 1_{\{i/2^n \leqslant f < (i+1)/2^n\}} + n 1_{\{f \geqslant n\}}.$$

Then

$$0 \leqslant f - f_n \leqslant \frac{1}{2^n} \text{ if } f < n;\ f_n = n \leqslant f \text{ otherwise.}$$

2. $f = f^+ - f^-$. $\square$

Thm 3 (Doob). $f : (E, \mathscr{F}_E) \mapsto (\mathbb{R}, \mathscr{B}(\mathbb{R}))$, *$g$ measurable* $(E, \mathscr{F}_E) \mapsto (F, \mathscr{F}_F)$. *If f is $\sigma(g)$-measurable, then $f = h \circ g$ for some measurable h.*

PROOF. **1**. $f = 1_A$, $A = g^{-1}(B) \in \sigma(g)$, $B \in \mathscr{F}_F$. Then $x \in A$ if and only if $g(x) \in B$, i.e.,

$$f = 1_A = 1_B \circ g.$$

2. f simple, $f = \sum_{i=1}^{n} c_i 1_{A_i}$, $c_i \in \mathbb{R}$, $A_i \in \sigma(g)$ disjoint. Let $A_i = g^{-1}(B_i)$, $B_i \in \mathscr{F}_F$, then

$$C_i = B_i \Big\backslash \Big(\bigcup\nolimits_{j<i} B_j\Big) \in \mathscr{F}_F \text{ disjoint}$$

and

$$f^{-1}(C_i) = A_i \Big\backslash \Big(\bigcup\nolimits_{j<i} A_j\Big) = A_i.$$

By step **1**,

$$f = \sum_{i=1}^{n} c_i 1_{A_i} = \sum_{i=1}^{n} c_i 1_{C_i} \circ g = \left(\sum_{i=1}^{n} c_i 1_{C_i}\right) \circ g \triangleq h \circ g.$$

3. $f \geqslant 0$ is $\sigma(g)$-measurable, there exist $\sigma(g)$-measurable simple f_n with $0 \leqslant f_n \uparrow f$. It follows $f_n = h_n \circ g$ for some h_n,

$$h \triangleq \sup_n h_n$$

is $\sigma(g)$-measurable,

$$f = \lim_n f_n = \sup_n (h_n \circ g) = \left(\sup_n h_n\right) \circ g = h \circ g.$$

4. f is $\sigma(g)$-measurable. f^+, f^- are $\sigma(g)$-measurable. Use **3**. $\square$

Thm 4. *$\mathscr{A}$ is a π-system, $\Omega \in \mathscr{A}$, $\mathcal{H}$ is a collection of real-valued functions. Suppose*
(1) *If $A \in \mathscr{A}$, then $1_A \in \mathcal{H}$*
(2) *If $f, g \in \mathcal{H}$, $c \in \mathbb{R}$, then $f+g$, $cg \in \mathcal{H}$*
(3) *If $f_n \in \mathcal{H}$, $0 \leqslant f_n \uparrow f$ with f bounded, then $f \in \mathcal{H}$*
Then

$$\{f : f \textit{ bounded } \sigma(\mathscr{A})\textit{-measurable}\} \subset \mathcal{H}$$

Proof. The system of sets

$$\mathscr{G} = \{A : 1_A \in \mathcal{H}\}$$

is a λ-system and $\mathscr{A} \subset \mathscr{G}$. Hence

$$\sigma(\mathscr{A}) \subset \mathscr{G}.$$

(2) implies that $\mathcal{H}$ contains all $\sigma(\mathscr{A})$-measurable simple functions, (3) implies that $\mathcal{H}$ contains all bounded $\sigma(\mathscr{A})$-measurable functions. $\square$

Thm 5. *$\mathscr{A}$ is a π-system, $\Omega \in \mathscr{A}$, $\mathcal{H}$ is a collection of real-valued functions. Suppose*

(1) *If $A \in \mathscr{A}$, then $1_A \in \mathcal{H}$*

(2) *If $f, g \in \mathcal{H}$, $a, b \geqslant 0$, then $af + bg \in \mathcal{H}$*

(3) *If $f, g \in \mathcal{H}$ are bounded, $f \geqslant g$, then $f - g \in \mathcal{H}$*

(4) *If $f_n \in \mathcal{H}$, $0 \leqslant f_n \uparrow f$, then $f \in \mathcal{H}$*

Then

$$\{f : f \text{ nonnegative } \sigma(\mathscr{A})\text{-measurable}\} \subset \mathcal{H}$$

2. 集函数与测度

2.1. 集函数.

$\mathscr{E}$ is a collection of subsets of E.

Def 2. *Set function,* $\mu : \mathscr{E} \mapsto \mathbb{R} \cup \{\pm\infty\}$.

Def 3. *Nonnegative set function,* $\mu : \mathscr{E} \mapsto \mathbb{R} \cup \{\infty\}$.

Def 4. μ *is finite if,* $\forall A \in \mathscr{E}$, $|\mu(A)| < \infty$.

Def 5. μ *is* σ*-finite on* $\mathscr{E}$ *if,* $\forall A \in \mathscr{E}$, *there exist* $\{A_n\} \subset \mathscr{E}$, $A = \bigcup_n A_n$ *with* $|\mu(A_n)| < \infty$.

Def 6. μ *is additive if,* $\forall A, B \in \mathscr{E}$, $AB = \varnothing$,

$$\mu(A + B) = \mu(A) + \mu(B).$$

Def 7. μ *is countably additive if,* $\forall A_i \in \mathscr{E}$, $i = 1, 2, \ldots$, *disjoint,*

$$\mu\left(\sum_i A_i\right) = \sum_i \mu(A_i).$$

Def 8. $\varnothing \in \mathscr{E}$. *$\mu$ is a measure on $\mathscr{E}$ if it is nonnegative, countably additive, $\mu(\varnothing) = 0$.*

E.g. 1. $(X, \mathscr{F})$ *measurable space, $x \in X$,*

$$\delta_x(A) = 1_A(x), \ \forall A \in \mathscr{F}.$$

$x_1, ..., x_n \in X,$

$$\mu(A) = \sum_i \delta_{x_i}(A), \ \forall A \in \mathscr{F}.$$

E.g. 2. *F real-valued nonnegative, non-decreasing, right continuous. Semi-ring on $\mathbb{R}$,*

$$\mathscr{A} = \{(a, b] : a, b, \in \mathbb{R}\}.$$

Then

$$\mu((a, b]) = F(b) - F(a)$$

defines a measure $\mathscr{A}$. It is unique on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$.

Proof. **1**. Additivity. $(a_i, b_i]$, $i = 1, ..., n$, disjoint, $(a, b] = \bigcup_i^n (a_i, b_i]$, then

$$\mu((a, b]) = \sum_{i=1}^{n} \mu((a_i, b_i]).$$

2. $(a_i, b_i]$, $i = 1, ...,$ disjoint, $\bigcup_i (a_i, b_i] \subset (a, b]$, then

$$\sum_{i=1}^{\infty} \mu((a_i, b_i]) \leqslant \mu((a, b]).$$

3. $(a_i, b_i]$, $i = 1, ..., n$, $(a, b] \subset \bigcup_i^n (a_i, b_i]$, then

$$\mu((a, b]) \leqslant \sum_{i=1}^{n} \mu((a_i, b_i]).$$

4. $(a_i, b_i]$, $i = 1, ...,$ disjoint, $\bigcup_i (a_i, b_i] = (a, b]$, then

$$\mu((a,b]) = \sum_{i=1}^{\infty} \mu((a_i, b_i]).$$

$\forall \varepsilon > 0$, there is $\delta_i > 0$,

$$F(b_i + \delta_i) - F(b_i) < \frac{\varepsilon}{2^i}.$$

$\forall \theta > 0$, $\{(a_i, b_i + \delta_i) : i\}$ is an open cover of $[a + \theta, b]$, there exists n_0

$$(a + \theta, b] \subset \bigcup_i^{n_0} (a_i, b_i + \delta_i].$$

By **3.**,

$$\begin{aligned}\mu((a+\theta,b]) &\leqslant \sum_{i=1}^{n_0}\mu((a_i,b_i+\delta_i]) \\ &= \sum_{i=1}^{n_0}(F(b_i+\delta_i)-F(b_i)) \\ &\leqslant \sum_{i=1}^{n_0}(F(b_i)-F(b_i))+\sum_{i=1}^{n_0}\frac{\varepsilon}{2^i} \\ &\leqslant \sum_{i=1}^{\infty}(F(b_i)-F(b_i))+\varepsilon.\end{aligned}$$

□

2.2. 半环上非负集函数. $\mathscr{E}$ is a collection of subsets of E, μ is a nonnegative set function on $\mathscr{E}$.

Def 9. *Monotonicity:* $\forall A \subset B \in \mathscr{E}$,

$$\mu(A) \leqslant \mu(B).$$

Def 10. *Countably subadditive:* $\forall A_i \in \mathscr{E}$, $i = 1, 2, ...,$ $\bigcup_{i=1}^{\infty} A_i \in \mathscr{E}$,

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \leqslant \sum_{i=1}^{\infty} \mu(A_i).$$

Def 11. *Continuity from below:* $A_i \in \mathscr{E}$, $A_i \uparrow A \in \mathscr{E}$,

$$\lim_n \mu(A_i) = \mu(A).$$

Def 12. *Continuity from above:* $A_i \in \mathscr{E}$, $A_i \downarrow A \in \mathscr{E}$, $\mu(A_1) < \infty$,

$$\lim_n \mu(A_i) = \mu(A).$$

REMARK 1. ***Note*** *finiteness is part of the defintion of continuity from above.*

$\mathscr{S}$ is a semi-ring on E, μ is a nonnegative set function on $\mathscr{S}$.

Suppose μ is **additive**.

1. $\mu(\varnothing) = 0, +\infty$.

PROOF. $\varnothing \in \mathscr{S}$. By additivity

$$\mu(\varnothing) = \sum_{i=1}^{n} \mu(\varnothing).$$

$\mu(\varnothing)$ equals 0, or ∞. $\square$

2. Monotonicity.

PROOF. $A, B \in \mathscr{S}$, $A \subset B$. There exist disjoint $C_1, ..., C_k \in \mathscr{S}$,

$$B \backslash A = \bigcup_{i=1}^{k} C_i.$$

$$B = A \cup (B \backslash A) = A \cup \left(\bigcup_{i=1}^{k} C_i \right).$$

By additivity

$$\mu(B) = \mu(A) + \sum_{i=1}^{k} \mu(C_i) \geqslant \mu(A).$$

$\square$

Suppose μ is **countably additive**.

3. Continuity from below.

Proof. $A_i \in \mathscr{S}$, $A_i \uparrow A \in \mathscr{S}$. There exist disjoint $C_{n,1}, ..., C_{n,k_n} \in \mathscr{S}$,

$$B_n \triangleq A_n \backslash A_{n-1} = \bigcup_{i=1}^{k_n} C_{n,i}.$$

$(A_0 = \varnothing)$

$$\begin{aligned}
\mu(A) &= \mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \mu\left(\bigcup_{n=1}^{\infty}\bigcup_{i=1}^{k_n} C_{n,i}\right) \\
&= \sum_{n=1}^{\infty}\sum_{i=1}^{k_n} \mu(C_{n,i}) = \lim_N \sum_{n=1}^{N}\sum_{i=1}^{k_n} \mu(C_{n,i}) \\
&= \lim_N \mu\left(\bigcup_{n=1}^{N}\bigcup_{i=1}^{k_n} C_{n,i}\right) = \lim_n \mu(A_n).
\end{aligned}$$

$\square$

4. Continuity from above.

PROOF. (**WRONG PROOF**)$A_i \in \mathscr{S}$, $A_i \downarrow A \in \mathscr{S}$, $\mu(A_1) < \infty$. Clearly

$$\mu\left(\bigcap_{n=1}^{\infty} A_n\right) \leqslant \mu(A_i) \leqslant \mu(A_1) < \infty.$$

$$\lim_n \mu(A_n) = \mu\left(\bigcap_{n=1}^{\infty} A_n\right)$$

$$\Longleftrightarrow$$

$$\mu(A_1) - \lim_n \mu(A_n) = \mu(A_1) - \mu\left(\bigcap_{n=1}^{\infty} A_n\right)$$

$$\Longleftrightarrow$$

$$\lim_n \mu(A_1 \backslash A_n) = \mu\left(A_1 \Big\backslash \bigcap_{n=1}^{\infty} A_n\right) = \mu\left(\bigcup_{n=1}^{\infty} (A_1 \backslash A_n)\right).$$

$\square$

5. Subadditivity.

Proof. Analogous to continuity from below. $\square$

2.3. 环上非负集函数.

Thm 6. *$\mathscr{R}$ is a ring. μ is nonnegative additive.*
(1) *μ countably additive*

$$\Longleftrightarrow$$

(2) *μ countably subadditive*

$$\Longleftrightarrow$$

(3) *μ continuity from below*

$$\Longrightarrow$$

(4) *μ continuity from above*

$$\Longrightarrow$$

(5) *μ continuity from above at $\varnothing$.*

If μ is finite, (5) implies (1).

PROOF. **1**. Already have: (1) $\Longrightarrow$ (2), (1) $\Longrightarrow$ (3), (1) $\Longrightarrow$ (4), (4) $\Longrightarrow$ (5).

2. (2) $\Longrightarrow$ (1). Suppose $A_i \in \mathscr{R}$, $i = 1, 2, ...$, disjoint, $\bigcup_{i=1}^{\infty} A_i \in \mathscr{R}$.

By countable subadditivity,

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \leqslant \sum_{i=1}^{\infty} \mu(A_i).$$

By monotonicity and additivity,

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \geqslant \mu\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mu(A_i),\ \forall n.$$

Sending $n \to \infty$,

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \geqslant \sum_{i=1}^{\infty} \mu(A_i).$$

3. (3) $\Longrightarrow$ (1). Suppose $A_i \in \mathscr{R}$, $i = 1, 2, ...,$ disjoint, $\bigcup_{i=1}^{\infty} A_i \in \mathscr{R}$.

Since

$$\bigcup_{i=1}^{n} A_i \uparrow \bigcup_{i=1}^{\infty} A_i,$$

by continuity from below,

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_n \mu\left(\bigcup_{i=1}^{n} A_i\right) = \lim_n \sum_{i=1}^{n} \mu(A_i) = \sum_{i=1}^{\infty} \mu(A_i).$$

4. (5) $\Longrightarrow$ (1). Suppose $A_i \in \mathscr{R}$, $i = 1, 2, ...,$ disjoint, $\bigcup_{i=1}^{\infty} A_i \in \mathscr{R}$.

Then, $\forall n$,

$$\bigcup_{i=1}^{n} A_i \in \mathscr{R} \text{ and } \bigcup_{i=n+1}^{\infty} A_i = \bigcup_{i=1}^{\infty} A_i \Big\backslash \bigcup_{i=1}^{n} A_i \in \mathscr{R}.$$

By additivity

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \mu\left(\bigcup_{i=1}^{n} A_i\right) + \mu\left(\bigcup_{i=n+1}^{\infty} A_i\right).$$

Since μ is finite

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) < \infty.$$

The continuity from above at $\varnothing$ yields,

$$\lim_{n} \mu\left(\bigcup_{i=n+1}^{\infty} A_i\right) = 0.$$

Hence

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_n \mu\left(\bigcup_{i=1}^{n} A_i\right) + \lim_n \mu\left(\bigcup_{i=n+1}^{\infty} A_i\right)$$
$$= \lim_n \sum_{i=1}^{n} \mu(A_i) = \sum_{i=1}^{\infty} \mu(A_i).$$

$\square$

3. Carathéodory's 延拓

3.1. 外测度.

Def 13. *μ^* is an outer measure on E if*
(1) $\mu^*(\varnothing) = 0$
(2) $\forall A, B \in 2^E$, *if* $A \subset B$, *then*

$$\mu^*(A) \leqslant \mu^*(B)$$

(3) *If* $A_i \in 2^E$,$i = 1, 2, ...,$

$$\mu^*\left(\bigcup_{i=1}^{\infty} A\right) \leqslant \sum_{i=1}^{\infty} \mu^*(A_i)$$

Thm 7. *Let $\mathscr{E}$ be a collection of sets on E, $\varnothing \in \mathscr{E}$. μ is a nonnegative set function on $\mathscr{E}$ with $\mu(\varnothing) = 0$. Define, $\forall A \in 2^E$,*

$$\mu^*(A) = \inf\left\{\sum_{i=1}^{\infty} \mu(A_i) : A_i \in \mathscr{E},\ A \subset \bigcup_{i=1}^{\infty} A_i\right\}.$$

Then $\mu^*(A)$ is an outer measure.

Proof. **1**. $\mu^*(\varnothing) = 0$ since $\varnothing \in \mathscr{E}$, $\varnothing \subset \bigcup\limits_{i=1}^{\infty} \varnothing$.

2. If $A \subset B$, $B \subset \bigcup\limits_{i=1}^{\infty} B_i$, then $A \subset \bigcup\limits_{i=1}^{\infty} B_i$, from the definition $\mu^*(A) \leqslant \mu^*(B)$.

3. Let $A_i \in 2^E, i = 1, 2, ..., \varepsilon > 0$. There are $A_{i,k} \in \mathscr{E}$, $A_i \subset \bigcup\limits_{k=1}^{\infty} A_{i,k}$,

$$\sum_{k=1}^{\infty} \mu(A_{i,k}) \leqslant \mu^*(A_i) + \frac{\varepsilon}{2^i},\ \forall i.$$

Since

$$\bigcup_{i=1}^{\infty} A_i \subset \bigcup_{i=1}^{\infty} \bigcup_{k=1}^{\infty} A_{i,k},$$

$$\mu^*\left(\bigcup_{i=1}^{\infty} A_i\right) \leqslant \sum_{i=1}^{\infty}\sum_{k=1}^{\infty} \mu(A_{i,k})$$
$$\leqslant \sum_{i=1}^{\infty}\left[\mu^*(A_i) + \frac{\varepsilon}{2^i}\right] \leqslant \sum_{i=1}^{\infty} \mu^*(A_i) + \varepsilon.$$

$\square$

Def 14. *μ^* is an outer measure on E. $A \in 2^E$ is μ^*-measurable if*

$$\mu^*(D) = \mu^*(D \cap A) + \mu^*(D \cap A^c),\ \forall D \in 2^E.$$

The class of μ^*-measurable sets is denoted by $\mathscr{F}_\mu^*$.

Def 15. *Let μ be a measure on a σ-field $\mathscr{F}$ of E, the measure space $(E, \mathscr{F}, \mu)$ is complete if*

$$A \in \mathscr{F},\ \mu(A) = 0 \Longrightarrow B \in \mathscr{F},\ \forall B \subset A.$$

Thm 8 (Carathéodory). *Let $\mathscr{E}$ be a collection of sets on E, $\varnothing \in \mathscr{E}$. μ is a nonnegative set function on $\mathscr{E}$ with $\mu(\varnothing) = 0$.*

(1) $\mathscr{F}_\mu^*$ is a σ-field.

(2) $\left(E, \mathscr{F}_\mu^*, \mu^*\right)$ is a complete measure space.

PROOF. **1**. Obviously, $E \in \mathscr{F}_\mu^*$ and $A^c \in \mathscr{F}_\mu^*$ if $A \in \mathscr{F}_\mu^*$.

2. If $A_1, A_2 \in \mathscr{F}_\mu^*$, then $A_1 \cup A_2$, $A_1 \cap A_2 \in \mathscr{F}_\mu^*$.

$\forall D \in 2^E$, we note

$$D \cap (A_1 \cup A_2) = (D \cap A_1) \cup (D \cap A_1^c \cap A_2).$$

Then

$$\begin{aligned}
&\mu^*(D \cap (A_1 \cup A_2)) + \mu^*(D \cap (A_1 \cup A_2)^c) \\
&\leqslant \mu^*(D \cap A_1) + \mu^*(D \cap A_1^c \cap A_2) + \mu^*(D \cap A_1^c \cap A_2^c) \text{ (subadditivity)} \\
&\leqslant \mu^*(D \cap A_1) + \mu^*(D \cap A_1^c) \ (A_2 \in \mathscr{F}_\mu^*) \\
&= \mu^*(D) \ (A_1 \in \mathscr{F}_\mu^*).
\end{aligned}$$

Hence

$$A_1 \cup A_2 \in \mathscr{F}^*_\mu.$$

It follows that

$$(A_1 \cap A_2)^c = A_1^c \cup A_2^c \in \mathscr{F}^*_\mu.$$

3. Finite additivity. If $A_1, ..., A_n \in \mathscr{F}^*_\mu$ disjoint, then $\forall D \in 2^E$,

$$\mu^*\left(D \cap \left(\bigcup_{i=1}^n A_i\right)\right) = \sum_{i=1}^n \mu^*(D \cap A_i).$$

Indeed, since $A_1 \in \mathscr{F}_\mu^*$,

$$\begin{aligned}
&\mu^*\left(D \cap \left(\bigcup_{i=1}^n A_i\right)\right) \\
&= \mu^*\left(D \cap \left(\bigcup_{i=1}^n A_i\right) \cap A_1\right) + \mu^*\left(D \cap \left(\bigcup_{i=1}^n A_i\right) \cap A_1^c\right) \\
&= \mu^*(D \cap A_1) + \mu^*\left(D \cap \left(\bigcup_{i=2}^n A_i\right)\right) = \cdots = \sum_{i=1}^n \mu^*(D \cap A_i)
\end{aligned}$$

4. If $A_1, A_2, ..., \in \mathscr{F}_\mu^*$, then $A \triangleq \bigcup_{i=1}^\infty A_i \in \mathscr{F}_\mu^*$.

We can assume that $A_1, A_2, ..., \in \mathscr{F}_\mu^*$ are disjoint. Indeed, by **1** and **2**, $B_i = A_i \Bigg\backslash \left(\bigcup_{j<i} A_j\right) \in \mathscr{F}_\mu^*$, are disjoint and $\bigcup_{i=1}^n B_i = \bigcup_{i=1}^n A_i$,

$\forall n$. Let

$$C_n = \bigcup_{i=1}^{n} A_i \in \mathscr{F}_\mu^*, \ \forall n.$$

Since $A_1, A_2, ...$ are disjoint, we can use **3** (the finite additivity). $\forall D \in 2^E$,

$$\begin{aligned}\mu^*(D) &= \mu^*(D \cap C_n) + \mu^*(D \cap C_n^c) \\ &= \sum_{i=1}^{n} \mu^*(D \cap C_i) + \mu^*(D \cap C_n^c) \\ &\geqslant \sum_{i=1}^{n} \mu^*(D \cap C_i) + \mu^*(D \cap A^c), \ \forall n.\end{aligned}$$

Let $n \to \infty$, note $A \subset \bigcup_{i=1}^{\infty} C_i$ and use subadditivity of outer measure

$$\mu^*(D) \geqslant \sum_{i=1}^{\infty} \mu^*(D \cap C_i) + \mu^*(D \cap A^c) \geqslant \mu^*(D \cap A) + \mu^*(D \cap A^c).$$

5. Countable additivity.

If $A_1, A_2, ..., \in \mathscr{F}_\mu^*$ are disjoint, use **3** and send $n \to \infty$,

$$\mu^*\left(\bigcup_{i=1}^{\infty} A_i\right) \geqslant \mu^*\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mu^*(A_i),\ \forall n.$$

The opposite inequality is subadditivity of outer measure.

6. Completeness. If $A \in \mathscr{F}_\mu^*$, $\mu^*(A) = 0$ and $B \subset A$, then $\mu^*(B) = 0$. $\forall D \in 2^E$,

$$\mu^*(D) \geqslant \mu^*(D \cap B^c) = \mu^*(D \cap B) + \mu^*(D \cap B^c).$$

So $B \in \mathscr{F}_\mu^*$. □

3.2. 域上测度的延拓.

Thm 9. *If μ is a measure on a field $\mathscr{A}$ with the generated outer measure μ^*. Then*

(1) $\mathscr{A} \subset \mathscr{F}_\mu^*$ *thus* $\sigma(\mathscr{A}) \subset \mathscr{F}_\mu^*$.
(2) μ^* *is an extension of* μ *to* $\sigma(\mathscr{A})$ *in the sense that*

$$\mu(A) = \mu^*(A), \ \forall A \in \mathscr{A}.$$

PROOF. **1**. Let $A \subset \mathscr{A}$. If $A_i \in \mathscr{A}$, $A \subset \bigcup_{i=1}^{\infty} A_i$, then

$$\mu(A) \leqslant \sum_{i=1}^{\infty} \mu(A_i). \tag{3.1}$$

Indeed,

$$\mu\left(A \cap \bigcup_{i=1}^{n} A_i\right) \leqslant \mu\left(\bigcup_{i=1}^{n} A_i\right) \leqslant \sum_{i=1}^{n} \mu(A_i).$$

Let $n \to \infty$ and use that μ is a measure to get (3.1). So

$$\mu(A) \leqslant \mu^*(A).$$

Since $A \subset \mathscr{A}$, $A_1 = A$, $A_2 = A_3 ... = \varnothing$ form a countable cover of A, so

$$\mu^*(A) \leqslant \mu(A).$$

2. Fix $A \subset \mathscr{A}$, will prove $A \in \mathscr{F}_\mu^*$. $\forall D \in 2^E$, it is enough to show that

$$\mu^*(D) \geqslant \mu^*(D \cap A) + \mu^*(D \cap A^c).$$

There is nothing to prove if $\mu^*(D) = \infty$, so we assume that $\mu^*(D) < \infty$. Then, $\forall \varepsilon > 0$, there exist $A_i \in \mathscr{A}$, $D \subset \bigcup\limits_{i=1}^{\infty} A_i$ so that

$$\sum_{i=1}^{\infty} \mu(A_i) \leqslant \mu^*(D) + \varepsilon.$$

Since $\mathscr{A}$ is a field,

$$A_i \cap A, A_i \cap A^c \in \mathscr{A}.$$

By **1** and the additivity of μ,

$$\begin{aligned}\mu(A_i) &= \mu(A_i \cap A) + \mu(A_i \cap A^c) \\ &= \mu^*(A_i \cap A) + \mu^*(A_i \cap A^c).\end{aligned}$$

Summing over i gives

$$\begin{aligned}\sum_{i=1}^{\infty} \mu(A_i) &= \sum_{i=1}^{\infty} \mu^*(A_i \cap A) + \sum_{i=1}^{\infty} \mu^*(A_i \cap A^c) \\ &\geqslant \mu^*(D \cap A) + \mu^*(D \cap A^c).\end{aligned}$$

So

$$\mu^*(D) + \varepsilon \geqslant \sum_{i=1}^{\infty} \mu(A_i) \geqslant \mu^*(D \cap A) + \mu^*(D \cap A^c).$$

$\square$

Thm 10 (Uniqueness)**.** *Let $\mathscr{P}$ be a π-system on E, μ and ν measures on $\sigma(\mathscr{P})$. Assume that*

(1) *μ and ν agree on $\mathscr{P}$.*

(2) *There are* $B_i \in \mathscr{P}$, $i = 1, 2, ...$, *disjoint so that* $\bigcup_{i=1}^{\infty} B_i = E$ *and* $\mu(B_i) < \infty$.

Then μ *and* ν *are equal on* $\sigma(\mathscr{P})$.

PROOF. **1**. Let $B \in \mathscr{P}$ have $\mu(B) < \infty$. Define

$$\mathscr{L} = \{A \in \sigma(\mathscr{P}) : \mu(A \cap B) = \nu(A \cap B)\}.$$

$\mathscr{L}$ is a λ-system (finiteness is needed to justify sets subtraction!), $\mathscr{P} \subset \mathscr{L}$. So

$$\sigma(\mathscr{P}) \subset \mathscr{L},$$

i.e.

$$\mu(A \cap B) = \nu(A \cap B), \ \forall A \in \sigma(\mathscr{P}).$$

2. $\forall A \in \sigma(\mathscr{P})$, use (2) to write it as disjoint union,

$$A = \bigcup_{i=1}^{\infty} (A \cap B_i), \ \mu(A \cap B_i) \leqslant \mu(B_i) < \infty.$$

Then, by **1**,

$$\mu(A) = \mu\left(\bigcup_{i=1}^{\infty}(A \cap B_i)\right) = \sum_{i=1}^{\infty} \mu(A \cap B_i)$$
$$= \sum_{i=1}^{\infty} \nu(A \cap B_i) = \nu\left(\bigcup_{i=1}^{\infty}(A \cap B_i)\right) = \nu(A).$$

□

▷ 2. *The condition Therem 10 (2) can be replaced with either one of the following:*
(2′) $\mathscr{P}$ *is a semi-ring,* $E \in \mathscr{P}$ *and* μ *is* σ*-finite on* $\mathscr{P}$.
(2″) *there are* $B_1, B_2, \ldots \in \mathscr{P}$, *so that* $B_i \uparrow E$ *and* $\mu(B_i) < \infty$.

3.3. 半环上测度的延拓.

Thm 11. *Let* μ *be a measure on the semi-ring* $\mathscr{S}$ *with the generated outer measure* μ^*. *Then*
(1) $\mathscr{S} \subset \mathscr{F}_\mu^*$ *thus* $\sigma(\mathscr{S}) \subset \mathscr{F}_\mu^*$.

(2) *μ^* is an extension of μ to $\sigma(\mathscr{S})$ in the sense that*

$$\mu(A) = \mu^*(A), \ \forall A \in \mathscr{S}. \tag{3.2}$$

(3) *Assume that there are $B_i \in \mathscr{S}$, $i = 1, 2, ...,$ disjoint so that $\bigcup_{i=1}^{n} B_i = E$ and $\mu(B_i) < \infty$, then the extension of μ to $\sigma(\mathscr{S})$ is unique.*

PROOF. Let $\bar{\mu}$ be the outer measure generated by μ.

1. $\bar{\mu}$ agrees with μ on $\mathscr{S}$.

The proof is identical to Theorem 9 (1).

2. Fix $A \subset \mathscr{S}$, will prove $A \in \mathscr{F}^*_\mu$.

The proof is identical to Theorem 9 (2). The difference is $A_i \cap A^c$ is replaced with disjoint union of sets in $\mathscr{S}$.

3. Uniqueness. Apply Theorem 10 to conclude. $\square$

3.4. Approximating $\mu^*|_{\mathscr{F}_\mu^*}$ by $\mu^*|_{\sigma(\mathscr{S})}$.

Thm 12. *Let μ be a measure on the semi-ring $\mathscr{S}$ with the generated outer measure μ^*. Suppose $E \in \mathscr{S}$.*

(1) $\forall A \in \mathscr{F}_\mu^*$, *there is $B \in \sigma(\mathscr{S})$ such that $A \subset B$ and*

$$\mu^*(A) = \mu^*(B).$$

(2) *If μ is σ-finite on $\mathscr{S}$, then $\forall A \in \mathscr{F}_\mu^*$, there is $B \in \sigma(\mathscr{S})$ such that $A \subset B$ and*

$$\mu^*(B \backslash A) = 0.$$

Proof.

1. There is nothing to prove if $\mu^*(A) = \infty$, we assume that $\mu^*(A) < \infty$. There are $B_{n,i} \in \mathscr{S}$, $A \subset \bigcup_{i=1}^{\infty} B_{n,i}$,

$$\sum_{i=1}^{\infty} \mu(B_{n,i}) < \mu^*(A) + \frac{1}{n}.$$

Set

$$B = \bigcap_{n=1}^{\infty} \bigcup_{i=1}^{\infty} B_{n,i}.$$

Then $A \subset B \in \sigma(\mathscr{S})$,

$$\mu^*(A) \leqslant \mu^*(B).$$

Moreover

$$\mu^*(B) \leqslant \mu^* \left(\bigcup_{i=1}^{\infty} B_{n,i} \right) \leqslant \sum_{i=1}^{\infty} \mu(B_{n,i}) \leqslant \mu^*(A) + \frac{1}{n}.$$

It follows that

$$\mu^*(B) \leqslant \mu^*(A).$$

2. If μ is *finite* on $\mathscr{S}$, then by **1**, $\forall A \in \mathscr{F}_\mu^*$, there is $B \in \sigma(\mathscr{S})$ such that $A \subset B$ and

$$\mu^*(A) = \mu^*(B).$$

Since μ^* is a measure on $\mathscr{F}_\mu^*$, this gives

$$\mu^*(B \backslash A) = 0.$$

The σ-finite case follows from similar argument as in step **3** of Theorem 11. $\square$

3.5. Approximating $\mu|_{\sigma(\mathscr{A})}$ by $\mu|_{\mathscr{A}}$.

Thm 13. *Let μ be a measure on the field $\mathscr{A}$ with the generated outer measure μ^*. For any $A \in \sigma(\mathscr{A})$ with $\mu^*(A) < \infty$, $\forall \varepsilon > 0$, there is $B \in \mathscr{A}$ such that $\mu^*(A \Delta B) < \varepsilon$.*

If, in the last Theorem, the measure μ is defined on $\sigma(\mathscr{A})$ and σ-finite on $\mathscr{A}$, then μ must equal μ^* on $\sigma(\mathscr{A})$ by uniqueness, we can use μ in place of μ^* in the conclusion.

Thm 14. *Let $\mathscr{A}$ be a field, μ a measure on $\sigma(\mathscr{A})$ and σ-finite on $\mathscr{A}$. For any $A \in \sigma(\mathscr{A})$ with $\mu(A) < \infty$, $\forall \varepsilon > 0$, there is $B \in \mathscr{A}$ such that $\mu(A \Delta B) < \varepsilon$.*

3.6. Completion of a measure space.

Thm 15. *Let $(X, \mathscr{F}, \mu)$ be a measure space,*

$$\bar{\mathscr{F}} \triangleq \{A \cup N : A \in \mathscr{F}, N \subset B \text{ for some } B \in \mathscr{F} \text{ with } \mu(B) = 0\}.$$

Define

$$\bar{\mu}(A \cup N) = \mu(A), \ \forall A \in \bar{\mathscr{F}}.$$

Then $\left(X, \bar{\mathscr{F}}, \bar{\mu}\right)$ *is a complete measure space.*

Clearly the Theorem says

$$\bar{\mu}(A) = \mu(A), \ \forall A \in \mathscr{F}.$$

PROOF. **1**. $\bar{\mathscr{F}}$ is a σ-field.

Suppose $A \cup N \in \bar{\mathscr{F}}$ where $A \in \mathscr{F}$, $N \subset B$, $B \in \mathscr{F}$ with $\mu(B) = 0$. Then

$$(A \cup N)^c = (A^c \cap B^c) \cup (B \cap A^c \cap N^c) \in \bar{\mathscr{F}}.$$

Suppose $A_i \cup N_i \in \bar{\mathscr{F}}$ where $A_i \in \mathscr{F}$, $N_i \subset B_i$, $B_i \in \mathscr{F}$ with $\mu(B_i) = 0$. Then

$$\bigcup_{i=1}^{\infty}(A_i \cup N_i) = \left(\bigcup_{i=1}^{\infty} A_i\right) \cup \left(\bigcup_{i=1}^{\infty} N_i\right) \in \bar{\mathscr{F}},$$

since

$$\bigcup_{i=1}^{\infty} N_i \subset \bigcup_{i=1}^{\infty} B_i \in \mathscr{F}$$

and

$$\mu\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} \mu(B_i) = 0.$$

2. The definition of $\bar{\mu}$ nonambiguous, i.e.

$$A_1 \cup N_1 = A_2 \cup N_2 \in \bar{\mathscr{F}} \Longrightarrow \bar{\mu}(A_1 \cup N_1) = \bar{\mu}(A_2 \cup N_2).$$

Here $N_i \subset B_i$ for some $B_i \in \mathscr{F}$ with $\mu(B_i) = 0$, $i = 1, 2$.

$$\bar{\mu}(A_1 \cup N_1) = \mu(A_1) = \mu(A_1 \cup B_1 \cup B_2) \geqslant \mu(A_2) = \bar{\mu}(A_2 \cup N_2).$$

By symmetry,

$$\bar{\mu}(A_1 \cup N_1) \leqslant \bar{\mu}(A_2 \cup N_2).$$

(In fact

$$A_1 \cup B_1 \cup B_2 = A_1 \cup N_1 \cup B_1 \cup B_2 = A_2 \cup N_2 \cup B_1 \cup B_2 = A_2 \cup B_1 \cup B_2$$

so

$$\mu(A_1 \cup B_1 \cup B_2) = \mu(A_2).$$

)

3. Countable additivity. Suppose $A_i \cup N_i \in \bar{\mathscr{F}}$ disjoint, where $A_i \in \mathscr{F}$, $N_i \subset B_i$, $B_i \in \mathscr{F}$ with $\mu(B_i) = 0$. Then

$$\bar{\mu}\left(\bigcup_{i=1}^{\infty}(A_i \cup N_i)\right) = \mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty}\mu(A_i) = \sum_{i=1}^{\infty}\bar{\mu}(A_i \cup N_i).$$

4. Completeness. Let $A \cup N \in \bar{\mathscr{F}}$, $N \subset B$, $B \in \mathscr{F}$ with $\mu(B) = 0$ and $\bar{\mu}(A \cup N)$, then

$$\mu(A \cup B) = \mu(A) = \bar{\mu}(A \cup N) = 0.$$

So for any $C \subset A \cup N$, $C \subset A \cup B$,

$$C = \varnothing \cup C \in \bar{\mathscr{F}}.$$

$\square$

Thm 16. *Suppose that μ is σ-finite on the semi-ring $\mathscr{S}$ with the generated outer measure μ^*. Then $\left(X, \mathscr{F}_\mu^*, \mu^*\right)$ is the completion of $(X, \sigma(\mathscr{S}), \mu^*)$.*

PROOF. Let

$\bar{\mathscr{F}} \triangleq \{A \cup N : A \in \sigma(\mathscr{S}), N \subset B$ for some $B \in \sigma(\mathscr{S})$ with $\mu(B) = 0\}$.

It is enough to show that

$$\mathscr{F}_\mu^* = \bar{\mathscr{F}}.$$

Since $\left(X, \mathscr{F}_\mu^*, \mu^*\right)$ is a complete measure space,

$$\bar{\mathscr{F}} \subset \mathscr{F}_\mu^*.$$

Let $A \in \mathscr{F}_\mu^*$, by Theorem 12 there exist $B, C \in \sigma(\mathscr{S})$ so that

$$A \subset B, \ \mu^*(B \backslash A) = 0; \ B \backslash A \subset C, \ \mu^*(C) = \mu^*(B \backslash A) = 0.$$

Writing

$$A = (B \cap C^c) \cup (A \cap C),$$

we get that $B \cap C^c \in \sigma(\mathscr{S})$, $(A \cap C) \subset C$, $\mu^*(C) = 0$, so $A \in \bar{\mathscr{F}}$. $\square$

4. 收敛

4.1. 可测函数的收敛. $(E, \mathscr{F}, \mu)$ a measure space, $f_n \in \mathscr{F}$, $i = 1, 2, ...$, $f \in \mathscr{F}$

Def 16. *Almost everywhere convergence,* $f_n \xrightarrow{a.e.} f$*:*

$$\mu\Big(\lim_n f_n \neq f\Big) = 0.$$

Def 17. *Convergence in measure,* $f_n \xrightarrow{\mu} f$*:* $\forall \varepsilon > 0$,

$$\lim_n \mu(|f_n - f| > \varepsilon) = 0.$$

Evidently

$$\begin{aligned} f_n \xrightarrow{a.e.} f &\iff \forall \varepsilon > 0,\ \mu\left(\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \{|f_m - f| > \varepsilon\}\right) = 0 \\ &\iff \forall \varepsilon > 0,\ \mu(\{|f_n - f| > \varepsilon\} \text{ i.o.}) = 0. \end{aligned}$$

Recall

$$x \in \limsup A_n \iff x \in A_n \text{ i.o.}$$

Thm 17. *If μ is finite, then*

$$f_n \xrightarrow{a.e.} f \implies f_n \xrightarrow{\mu} f.$$

Proof. Indeed,

$$\mu(|f_n - f| > \varepsilon) \leqslant \mu\left(\bigcup_{m=n}^{\infty} \{|f_m - f| > \varepsilon\}\right), \ \forall n.$$

Let $n \to \infty$ and use continuity from above (requires finiteness of μ)

$$\begin{aligned} \limsup_n \mu(|f_n - f| > \varepsilon) &\leqslant \lim_n \mu\left(\bigcup_{m=n}^{\infty} \{|f_m - f| > \varepsilon\}\right) \\ &= \mu\left(\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \{|f_m - f| > \varepsilon\}\right) = 0. \end{aligned}$$

(or use

$$\limsup_n \mu(A_n) \leqslant \mu\left(\limsup_n A_n\right).$$

) ∎

Def 18. *Almost uniform convergence,* $f_n \xrightarrow{a.u.} f$*:* $\forall \varepsilon > 0$*, there is* $A_\varepsilon \in \mathscr{F}$ *so that* $\mu(A_\varepsilon) < \varepsilon$*,*

$$\lim_n \sup_{x \notin A_\varepsilon} |f_n - f| = 0.$$

Compare with Egoroff's Theorem on *finite* measure!

Thm 18. $f_n \xrightarrow{a.u.} f$ *if and only if* $\forall \varepsilon > 0$,

$$\lim_n \mu\left(\bigcup_{m=n}^{\infty} \{|f_m - f| > \varepsilon\}\right) = 0.$$

Proof. **1**. " $\Longrightarrow$ ". $\forall \varepsilon > 0$, there is A_ε so that $\mu(A_\varepsilon) < \varepsilon$ and

$$\lim_m \sup_{x \notin A_\varepsilon} |f_m - f| = 0.$$

So, $\forall \varepsilon' > 0$, there is $n_0 \in \mathbb{N}$ such that

$$\sup_{x \notin A_\varepsilon} |f_m - f| \leqslant \varepsilon', \ \forall m \geqslant n_0.$$

This translates to

$$\bigcup_{m=n_0}^{\infty} \{|f_m - f| > \varepsilon'\} \subset A_\varepsilon.$$

Therefore

$$\mu\left(\bigcup_{m=n_0}^{\infty} \{|f_m - f| > \varepsilon'\}\right) \leqslant \mu(A_\varepsilon) < \varepsilon.$$

2. " $\Longleftarrow$ ". $\forall \varepsilon > 0$ and $k \in \mathbb{N}$, there is $n_{\varepsilon,k} \in \mathbb{N}$ such that

$$\mu\left(\bigcup_{m=n_{\varepsilon,k}}^{\infty} \left\{|f_m - f| > \frac{1}{k}\right\}\right) < \frac{\varepsilon}{2^k},\ \forall m \geqslant n_{\varepsilon,k}.$$

Denote (the set of all possible divergence points! measurable!)

$$A_\varepsilon = \bigcup_{k=1}^{\infty} \bigcup_{m=n_{\varepsilon,k}}^{\infty} \left\{|f_m - f| > \frac{1}{k}\right\}.$$

Then $\mu(A_\varepsilon) < \varepsilon$ and for any $x \notin A_\varepsilon$, we have $\forall k$,

$$|f_m - f| \leqslant \frac{1}{k},\ \forall m > n_{\varepsilon,k}.$$

$\square$

We have proved:

Thm 19. (1)

$$f_n \xrightarrow{a.u.} f \implies f_n \xrightarrow{a.e.} f \ and\ f_n \xrightarrow{\mu} f$$

(2) *If μ is finite, then*

$$f_n \xrightarrow{a.u.} f \iff f_n \xrightarrow{a.e.} f \implies f_n \xrightarrow{\mu} f$$

E.g. 3.

$$f_n(x) = \begin{cases} 1, & x \in (0, 1/n), \\ 0, & x \in [1/n, 1]. \end{cases}$$

E.g. 4.

$$f_n(x) = x^n, x \in [0, 1]$$

$\triangleright$ 3. *Let* $f = 0$ *and* $f_n = 1_{A_n}$. *Then* $f_n \xrightarrow{\mu} f$ *is equivalent to* $\mu(A_n) \to 0$ *and* $\left(\lim_n f_n \neq f\right) = (A_n \ i.o.)$.

Any sequence $\{A_n\}$ so that $\mu(A_n) \to 0$ but $\mu(A_n \text{ i.o.}) > 0$ gives an exmple that $f_n \xrightarrow{\mu} f \not\Rightarrow f_n \xrightarrow{a.e.} f$. It is enough to have $\mu(A_n) \to 0$ and

$$\sum_{i=1}^{\infty} 1_{A_n}(x) = \infty, \ \sum_{i=1}^{\infty} 1_{A_n^c}(x) = \infty.$$

E.g. 5. *For each* $n = 1, 2, ...$ *there is a unique decomposition* $n = k(k-1)/2 + i$ *with* $k = 1, 2, ...,\ i = 1, 2, ..., k$.

$$f_n(x) = \begin{cases} 1, & x \in (((i-1)/k, i/k]), \\ 0, & \textit{otherwise.} \end{cases}$$

E.g. 6. *Consider*

$$A_k^i = \left[\frac{i-1}{k}, \frac{i}{k}\right],\ h_k^i(x) = 1_{A_k^i}(x),\ i = 1, ..., k.$$

Let f_n be the sequence

$$\{h_1^1; h_2^1, h_2^2; h_3^1, h_3^2; h_3^3; ...\}$$

Thm 20. *$f_n \xrightarrow{\mu} f \iff$ for any subsequence there is a further subsequence $f_{n_k} \xrightarrow{a.u.} f$.*

PROOF. " $\Longrightarrow$ ". Since any subsequence of f_n converges in measure to f, it is enough to show there is a subsequence $f_{n_k} \xrightarrow{a.u.} f$. To see this, for any $k > 0$, by definition of convergence in measure, we can choose $n_k > n_{k-1}$ so that

$$\mu\left(|f_{n_k} - f| > \frac{1}{k}\right) \leqslant \frac{1}{2^k}.$$

Then

$$\mu\left(\bigcup_{k=m}^{\infty} |f_{n_k} - f| > \frac{1}{k}\right) \leqslant \sum_{k=m}^{\infty} \frac{1}{2^k} = \frac{1}{2^{m-1}}.$$

$\forall \varepsilon > 0$, for large m,

$$\bigcup_{k=m}^{\infty} \{|f_{n_k} - f| > \varepsilon\} \subset \bigcup_{k=m}^{\infty} \left\{|f_{n_k} - f| > \frac{1}{k}\right\}.$$

So

$$\lim_{m} \mu\left(\bigcup_{k=m}^{\infty} |f_{n_k} - f| > \varepsilon\right) \leqslant \lim_{m} \mu\left(\bigcup_{k=m}^{\infty} |f_{n_k} - f| > \frac{1}{k}\right) = 0.$$

" $\Longleftarrow$ " Suppose $f_n \xrightarrow{\mu} f$ does not hold, i.e. there are $n_k \to \infty$, $\varepsilon_0 > 0$, $\delta_0 > 0$ so that

$$\mu(|f_{n_k} - f| > \varepsilon_0) > \delta_0.$$

Then

$$\liminf_{m} \mu\left(\bigcup_{k=m}^{\infty} |f_{n_k} - f| > \varepsilon_0\right) \geqslant \delta_0,$$

Contradicting Theorem 18. $\square$

Theorem 19 and Theorem 20 indicate that if $f_n \xrightarrow{\mu} f$, then there is a subsequence $f_{n_k} \xrightarrow{a.e.} f$.

4.2. 随机变量的分布函数.

Def 19. $(\Omega, \mathscr{F}, P)$ *is a probability space if P is a nonnegative measure on the σ-field $\mathscr{F}$ with $P(\Omega) = 1$.*

Def 20. *A random variable (r.v.) X on $(\Omega, \mathscr{F}, P)$ is a real-valued mapping, $X : \omega \in \Omega \mapsto X(\omega) \in \mathbb{R}$.*

Def 21. *The distribution function of a r.v. X is*

$$F(x) = P(X \leqslant x).$$

Denoted by $X \sim F$.

Thm 21. *Any distribution function F has the following properties.*
(1) *non-decreasing, $F(-\infty) = 0$ and $F(\infty) = 1$*
(2) *right continuity: $\lim_{y \downarrow x} F(y) = F(x)$.*
(3) *left limit exists: $F(x-) = \lim_{y \uparrow x} F(y) = P(X < x)$.*

(4) $P(X = x) = F(x) - F(x-)$.

The **inverse of the distrbution function** F is defined as below. $\forall z \in (0, 1)$,

$$F^{-1}(z) = \inf\{x \in \mathbb{R} : F(x) \geqslant z\}. \tag{4.1}$$

$\triangleright$ 4. *Also equivalently defined as,*

$$F^{-1}(z) = \sup\{x \in \mathbb{R} : F(x) < z\}. \tag{4.2}$$

LEMMA 22. *F^{-1} has the properties,*
(1) *F^{-1} is real-valued non-decreasing.*
(2) *F^{-1} is left-continuous and has right limit.*
(3) $F^{-1}(F(x)) \leqslant x$, $F\big(F^{-1}(z)\big) \geqslant z$.
(4) $F^{-1}(z) \leqslant x$ *iff* $F(x) \geqslant z$.

PROOF. Exercise. $\square$

Thm 23. *If F satisfies* (1)(2)(3) *of Theorem 21, there is a r.v. X with distribution F.*

PROOF. Let $\Omega = (0,1)$, $\mathscr{F} = \mathscr{B}_{(0,1)}$ (i.e. $(0,1) \cap \mathscr{B}_{\mathbb{R}}$), $P =$ Lebesgue measure. Define

$$X(\omega) = F^{-1}(\omega).$$

Then X is $\mathscr{F}$-measurable (check this!) and

$$\begin{aligned} P(\omega : X(\omega) \leqslant x) &= P(\omega : F(x) \geqslant \omega) \\ &= \text{Lebesgue measure of } (0, F(x)) = F(x). \end{aligned}$$

So X is a r.v. with distribution function F. □

▷ 5. *Another construction of a r.v. X with distribution F is to take $\Omega = (\mathbb{R}, \mathscr{B})$, $P =$ the Lebesgue measure induced by F and consider the coordinate map $X(\omega) = \omega$.*

4.3. 随机变量的收敛. Probability space $(\Omega, \mathscr{F}, P)$, r.v. X_n, X,

$$X_n \xrightarrow{a.s.} X \iff P(X_n = X) = 1.$$

$$X_n \xrightarrow{P} X \iff \forall \varepsilon > 0, \ \lim_n P(|X_n - X| > \varepsilon) = 0.$$

Def 22. $X_n \sim F_n$, $X \sim F$. *Convergence in distribution (weak convergence): $F_n(x) \to F(x)$ for all x where F is continuous, written $X_n \xrightarrow{d} X$.*

Thm 24. $X_n \sim F_n$, $X \sim F$.

$$X_n \xrightarrow{a.s.} X \implies X_n \xrightarrow{P} X \implies X_n \xrightarrow{d} X.$$

Proof. **1**. The first implication is a special case of Theorem 17.

2. Check the second implication. $\forall \varepsilon$, $x \in \mathbb{R}, n \in \mathbb{N}$,

$$\begin{aligned}
&P(X \leqslant x - \varepsilon) - P(|X_n - X| > \varepsilon) \\
&\leqslant P(X_n \leqslant x) \\
&\leqslant P(X_n \leqslant x, |X_n - X| \leqslant \varepsilon) + P(X_n \leqslant x, |X_n - X| > \varepsilon) \\
&\leqslant P(X \leqslant x + \varepsilon) + P(|X_n - X| > \varepsilon).
\end{aligned}$$

So $n \to \infty$, $\varepsilon \to 0$ yield

$$F(x-) \leqslant \liminf_n P(X_n \leqslant x) \leqslant \limsup_n P(X_n \leqslant x) \leqslant F(x).$$

$\square$

LEMMA 25. $F_n \xrightarrow{w} F \iff F_n^{-1} \xrightarrow{w} F^{-1}$.

PROOF OF " $\Longrightarrow$ ". Construct r.v.s' $X_n \sim F_n$, $X \sim F$ as Theorem 23. Fix any ω.

1. Choose any $\varepsilon > 0$ so that F is continuous at $X(\omega) - \varepsilon$ (the discontinuities of F are at most countable, ε can be arbitrarily small). By the definition (the infimum!) of $X(\omega)$,

$$F(X(\omega) - \varepsilon) < \omega.$$

Then, for large n,

$$F_n(X(\omega) - \varepsilon) < \omega.$$

so (note the above inequality is strict)

$$X(\omega) - \varepsilon \leqslant X_n(\omega).$$

Hence

$$X(\omega) \leqslant \liminf_n X_n(\omega).$$

2. To see the opposite. Choose any $\varepsilon, \delta > 0$ so that X is continuous at ω and F is continuous at $X(\omega)+\varepsilon$, then by Lemma 22

$$F(X(\omega+\delta)+\varepsilon) \geqslant F(X(\omega+\delta)) \geqslant \omega+\delta > \omega.$$

For large n ($\delta > 0$),

$$F_n(X(\omega+\delta)+\varepsilon) \geqslant \omega.$$

By Lemma 22 again,

$$X(\omega+\delta)+\varepsilon \geqslant X_n(F_n(X(\omega+\delta)+\varepsilon)) \geqslant X_n(\omega).$$

Let $n \to \infty$, $\varepsilon \to 0$, $\delta \to 0$ (continuity at ω),

$$X(\omega) \geqslant \limsup_n X_n(\omega).$$

$\square$

Thm 26 (Skorohod)**.** $X_n \sim F_n$, $X \sim F$. *Suppose* $X_n \xrightarrow{d} X$. *There exist r.v.* $\bar{X}_n$, $\bar{X}$ *on a common probability space so that* $\bar{X}_n \overset{d}{=} X_n$, $\bar{X} \overset{d}{=} X$, $\bar{X}_n \xrightarrow{a.s.} \bar{X}$.

PROOF. Let $\Omega = (0,1)$, $\mathscr{F} = \mathscr{B}_{(0,1)}$, $P =$ Lebesgue measure. By Theorem 23 there exist r.v. on $(\Omega, \mathscr{F}, P)$ so that $\bar{X}_n \sim F_n$, $\bar{X} \sim F$. Lemma 25 then says $F_n^{-1} \xrightarrow{w} F^{-1}$. Since the discontinuity set of F^{-1} is countable, $F_n^{-1}(\omega) \to F^{-1}(\omega)$ for almost all $\omega \in \Omega$, i.e. $\bar{X}_n(\omega) \xrightarrow{a.s.} \bar{X}(\omega)$. $\square$

5. 积分

5.1. 非负可测函数积分. $(E,\mathscr{F},\mu)$ a measure space, $f\in\mathscr{F}$ with values in $[0,\infty]$,. A *finite (measurable) partition* of E is a finite collection of $\mathscr{F}$-measurable sets $\{A_i : i=1,...,m\}$ with $\bigcup_{i=1}^{m} A_i = E$.

$$\int f d\mu \triangleq \sup_{\text{finite partitions}} \sum_i \left[\inf_{x\in A_i} f(x)\right]\mu(A_i). \tag{5.1}$$

Convention: $0\cdot\infty = 0$.

▷ 6. *Consider*

$$\int f d\mu \triangleq \inf_{\textit{finite partitions}} \sum_i \left[\sup_{x\in A_i} f(x)\right]\mu(A_i). \tag{5.2}$$

Is (5.2) *a good definition of integration?*

Properties: f, $g\in\mathscr{F}$ *nonnegative.*

(1) If $f=0$, μ-*a.e.*, then $\int f d\mu = 0$.

(2) If $\mu(f > 0) > 0$, then $\displaystyle\int f d\mu > 0$.

(3) If $\displaystyle\int f d\mu < \infty$, then $f < \infty, \mu$-$a.e.$

(4) If $f \leqslant g$, μ-$a.e.$, then $\displaystyle\int f d\mu \leqslant \int g d\mu$.

(5) If $f = g$, μ-$a.e.$, then $\displaystyle\int f d\mu = \int g d\mu$.

Thm 27 (Monotone convergence Theorem)**.** *If* $0 \leqslant f_n \uparrow f$, μ-$a.e.$, *then* $0 \leqslant \displaystyle\int f_n d\mu \uparrow \int f d\mu$.

PROOF. **1**. First prove it under the assumption that

$$0 \leqslant f_n(x) \uparrow f(x), \ \forall x.$$

Integration is monotonic, so $\displaystyle\int f_n d\mu \leqslant \int f d\mu$. It remains to show

$$\lim_n \int f_n d\mu \geqslant \int f d\mu \tag{5.3}$$

or

$$\lim_n \int f_n d\mu \geqslant S = \sum_{i=1}^{m} c_i \mu(A_i)$$

for any finite measurable partition $\{A_i : i = 1, ..., m\}$ and $c_i = \inf_{A_i} f$.

For such a partition, assume that the sum S, c_i and $\mu(A_i)$ are all finite. Fix $\alpha < 1$, define

$$A_{i,n} = \{x \in A_i : f_n(x) > \alpha c_i\}.$$

Since $f_n \uparrow f$, $A_{i,n} \uparrow A_i$. Consider the *measurable* partition

$$\{A_{i,n} : i = 1, ..., m\} \cup \left\{ \left(\bigcup_{i=1}^{m} A_{i,n} \right)^c \right\}.$$

Then

$$\int f_n d\mu \geqslant \sum_{i=1}^{m} \alpha c_i \mu(A_{i,n}).$$

Let $n \to \infty$ and use continuity from below,

$$\lim_n \int f_n d\mu \geqslant \sum_{i=1}^m \alpha c_i \mu(A_i).$$

Finally let $\alpha \to 1$, (5.3) is proved.

Now suppose S is finite but not all of c_i, $\mu(A_i)$. Then $c_i\mu(A_i)$, $i = 1, ..., m$ are finite. c_i or $\mu(A_i)$ may be infinity, but then $c_i\mu(A_i)$ must be zero. Use the adjusted parition $\{A_i : c_i\mu(A_i) > 0\} \cup \{\text{complement}\}$.

Lastly suppose S is infinite. Then there is some i_0, $c_{i_0}\mu(A_{i_0}) = \infty$, i.e., $c_{i_0} > 0$, $\mu(A_{i_0}) > 0$ and at least one of them is ∞. In this case

$$\int f d\mu = \infty.$$

To prove (5.3), let a, b satisfy

$$0 < a < c_{i_0} \leqslant \infty,\ 0 < b < \mu(A_{i_0}) \leqslant \infty.$$

Define

$$A_{i_0,n} = \{x \in A_{i_0} : f_n(x) > a\}.$$

Since $f_n \uparrow f$, $A_{i_0,n} \uparrow A_{i_0}$ and $\mu(A_{i_0,n}) > b$ for n larger than some $n_{a,b}$. For the partition $\{A_{i_0,n}, A_{i_0,n}^c\}$, we have

$$\int f_n d\mu \geqslant a\mu(A_{i_0,n}) > ab, \forall n > n_{a,b}.$$

Let $a \to \infty$ if $c_{i_0} = \infty$, $b \to \infty$ if $\mu(A_{i_0,n}) = \infty$, we get

$$\lim_n \int f_n d\mu = \infty.$$

2. If $0 \leqslant f_n \uparrow f$ on A with $\mu(A^c) = 0$, then $0 \leqslant f_n 1_A \uparrow f 1_A$ holds everywhere. Then apply step **1**. $\square$

5.2. 可测函数积分. $f \in \mathscr{F}$ with values in $[-\infty, \infty]$,

$$\int f d\mu \triangleq \int f^+ d\mu - \int f^- d\mu.$$

f is said to be integrable if $\displaystyle\int f^+ d\mu, \int f^- d\mu$ are finite. So f integrable iff $|f|$ integrable.

Properties: f, $g \in \mathscr{F}$ integrable.

(1) If $f \leqslant g$, μ-*a.e.*, then $\displaystyle\int f d\mu \leqslant \int g d\mu$.

(2) If $\alpha, \beta \in \mathbb{R}$, then $\alpha f + \beta g$ is integrable,

$$\int (\alpha f + \beta g) d\mu = \alpha \int f d\mu + \beta \int g d\mu.$$

E.g. 7. *Let* $E = \{1, 2, 3, ...\}$, $\mathscr{F} = \{$*all subsets of* $E\}$, $\mu =$ *counting measure. A function on* E *is a sequence* $x_1, x_2, ...$ *Any function is* $\mathscr{F}$*-measurable.* $\{x_k : k = 1, 2, ...\}$ *is* μ*-integrable if and only if* $\displaystyle\sum_{k=1}^{\infty} |x_k|$ *converges. When* μ*-integrable,*

$$\sum_{k=1}^{\infty} |x_k| = \sum_{k=1}^{\infty} x_k^+ - \sum_{k=1}^{\infty} x_k^-.$$

The function $x_k = (-1)^{k+1}/k$, $k = 1, 2, ...$ *is not* μ*-integrable, although*

$$\lim_{m} \sum_{k=1}^{m} (-1)^{k+1} \frac{1}{k} = \ln 2.$$

Thm 28 (Fatou's lemma)**.** *Given f_n measurable.*
(1) *If g integrable, $f_n \geqslant g$, μ-a.e, then $\liminf_n f_n$ is integrable and*

$$\int \liminf_n f_n d\mu \leqslant \liminf_n \int f_n d\mu.$$

(1) *If g integrable, $f_n \leqslant g$, μ-a.e, then $\limsup_n f_n$ is integrable and*

$$\limsup_n \int f_n d\mu \leqslant \int \limsup_n f_n d\mu.$$

Thm 29 (Lebesgue's dominated convergence theorem)**.** *Given g nonnegative integrable, $|f_n| \leqslant g$, μ-a.e.. If $f_n \xrightarrow{a.e.} f$ or $f_n \xrightarrow{\mu} f$, then*

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

The following is a generalized dominated convergence theorem.

Thm 30. *Given g_n nonnegative integrable, $|f_n| \leqslant g_n$, μ-a.e. with $g_n \xrightarrow{a.e.} g$ and $\int g_n d\mu \longrightarrow \int g d\mu$. If $f_n \xrightarrow{a.e.} f$ or $f_n \xrightarrow{\mu} f$, then*

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

E.g. 8 (Weierstrass M-test)**.** *If $|x_{n,m}| \leqslant M_m$, $\sum_{m=1}^{\infty} M_m < \infty$, $\lim_n x_{n,m} = x_m$ for each m. Then*

$$\lim_n \sum_{m=1}^{\infty} x_{n,m} = \sum_{m=1}^{\infty} x_m.$$

E.g. 9 (Bounded convergence theorem)**.** *Suppose μ is finite, $M > 0$. $|f_n| \leqslant M$, μ-a.e.. If $f_n \xrightarrow{a.e.} f$ or $f_n \xrightarrow{\mu} f$, then*

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

E.g. 10. *If* $f_n \geqslant 0$ *or* $\sum_{n=1}^{\infty} \int |f_n| d\mu < \infty$, *then*

$$\int \sum_{n=1}^{\infty} f_n d\mu = \sum_{n=1}^{\infty} \int f_n d\mu.$$

From this we get

E.g. 11. *If* $x_{n,m} \geqslant 0$ *or* $\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} |x_{n,m}| < \infty$, *then*

$$\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} x_{n,m} = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} x_{n,m}.$$

5.3. Change of variables. $(E_1, \mathscr{F}_1)$, $(E_2, \mathscr{F}_2)$ are measurable spaces, μ is a measure on $\mathscr{F}_1$. T is measurable mapping from $(E_1, \mathscr{F}_1)$ to $(E_2, \mathscr{F}_2)$. Define

$$\nu(B) = \mu\big(T^{-1}(B)\big), \ \forall B \in \mathscr{F}_2. \tag{5.4}$$

Then $\nu(B)$ is a measure on $\mathscr{F}_2$ and for any $f \in \mathscr{F}_2$,

$$\int_{E_2} f d\nu = \int_{E_1} f \circ T d\mu. \tag{5.5}$$

Note if $f = 1_B$, then $f \circ T(x) = 1_B(T(x)) = 1_{T^{-1}(B)}(x)$, since $T(x) \in B$ iff $x \in T^{-1}(B)$. So in this case (5.5) reduces to (5.4).

6. L_p 空间

6.1. Inequlities.

LEMMA 31 (Jensen's inequality). *Let $(\Omega, \mathscr{F}, \mu)$ be a measure space with $\mu(\Omega) = 1$, X a μ-integrable function on Ω, φ convex on $\mathbb{R}$. Then*

$$\varphi\Big(\int_\Omega X d\mu\Big) \leqslant \int_\Omega \varphi(X) d\mu. \tag{6.1}$$

Equality holds iff φ is linear on some convex set $A \subset \mathbb{R}$ with $\mu\big(X^{-1}A\big) = 1$.

PROOF. Denote by μ_X the induced measure of X on $\mathbb{R}$ (ref section 5.3), then (6.1) is equivalent to

$$\varphi\left(\int_{\mathbb{R}} x d\mu_X\right) \leqslant \int_{\mathbb{R}} \varphi(x) d\mu_X \tag{6.2}$$

(Apply (5.5) with $f(x) = x$, $T = X$). It is enough to prove (6.2).

1. Denote $\bar{x} = \displaystyle\int_{\mathbb{R}} x d\mu_X$. Since φ is convex, there is a supporting line $L(x) = ax + b$ through $\bar{x}$, i.e. $L(\bar{x}) = \varphi(\bar{x})$ and

$$L(x) \leqslant \varphi(x), \ \forall x.$$

Then

$$\int_{\mathbb{R}} L(x) d\mu_X \leqslant \int_{\mathbb{R}} \varphi(x) d\mu_X. \tag{6.3}$$

The LHS equals $\varphi\left(\displaystyle\int_{\mathbb{R}} x d\mu_X\right)$, hence (6.2) follows.

2. Suppose the equality in (6.2) holds, then by the above computation

$$\int_{\mathbb{R}} [\varphi(x) - L(x)] d\mu_X = 0.$$

The integrand is nonnegative, so the measurable set

$$A = \{x \in \mathbb{R} : \varphi(x) - L(x) = 0\}$$

has full measure, i.e. $\mu_X(A) = 1$. Moreover the set A is convex (verify directly!). On the other hand, if φ is linear on some convex $A \subset \mathbb{R}$ with $\mu\big(X^{-1}A\big) = 1$, then $\mu_X(A) = 1$,

$$\int_{\mathbb{R}} L(x) d\mu_X = \int_A L(x) d\mu_X, \quad \int_{\mathbb{R}} \varphi(x) d\mu_X = \int_A \varphi(x) d\mu_X.$$

Hence by (6.3),

$$\int_A [\varphi(X) - L(X)] d\mu \geqslant 0.$$

But the integrand $\varphi - L$ is nonnegative and linear on A. Since $A \subset \mathbb{R}$ is convex, it must be an interval. So the above integral is zero, hence the equality of (6.2) holds. $\square$

Notice that Lemma 31 does not require $\varphi(X)$ to be μ-integrable. From (6.3) it is clear that either $\int_\Omega \varphi(X)d\mu$ exists or equals infinity, in the latter case (6.1) trivially holds.

LEMMA 32. $a, b \in \mathbb{R}$, $1 \leqslant p < \infty$,

$$|a+b|^p \leqslant 2^{p-1}(|a|^p + |b|^p).$$

PROOF. Apply Jensen's inequality with $\varphi(x) = |x|^p$,

$$\left|\frac{a+b}{2}\right|^p \leqslant \frac{|a|^p + |b|^p}{2}.$$

$\square$

LEMMA 33 (Young's inequality). $a, b \geqslant 0$, $1 < p, q < \infty$, $\dfrac{1}{p} + \dfrac{1}{q} = 1$,

$$a^{1/p} b^{1/q} \leqslant \frac{a}{p} + \frac{b}{q}.$$

Equal iff $a = b$.

Proof. The inequality holds if $ab = 0$. In this case equality holds iff $a = b = 0$. Now suppose $ab > 0$. Apply Jensen's inequality with $\varphi(x) = -\ln x$,

$$-\ln\left(\frac{a}{p} + \frac{b}{q}\right) \leqslant -\frac{1}{p}\ln a - \frac{1}{q}\ln b.$$

Since φ is strictly convex (can touch a linear function at exactly one point), equality holds iff $a = b$. $\square$

$(E, \mathscr{F}, \mu)$ is a measure space in the following definitions.

Def 23. $p = 1$, *let*

$$L_1 \triangleq \{f \in \mathscr{F} : |f| \text{ is } \mu\text{-integrable}\}$$

and

$$\|f\|_1 = \|f\|_{L_1} = \int |f| d\mu.$$

Def 24. $1 < p < \infty$, *let*

$$L_p \triangleq \{f \in \mathscr{F} : |f|^p \in L_1\}$$

and

$$\|f\|_p = \|f\|_{L_p} = \left(\int |f|^p d\mu\right)^{1/p}.$$

Def 25. $p = \infty$, *let*

$$L_\infty \triangleq \{f \in \mathscr{F} : \textit{there is } C > 0 \textit{ such that } |f| \leqslant C, \ \textit{a.e.}\}$$

and

$$\|f\|_\infty = \|f\|_{L_\infty} = \inf\{C : |f| \leqslant C, \ \textit{a.e.}\}.$$

We could have written $L_p(\mu)$ to emphasize the dependence of the spaces L_p on the measure μ. But, when no ambiguity arises from the contexts, we will simply drop μ from the notation.

Thm 34 (Hölder inequality)**.** $1 \leqslant p, \ q \leqslant \infty, \ \dfrac{1}{p} + \dfrac{1}{q} = 1, \ f \in L_p,$ $g \in L_q$, *then* $fg \in L_1$ *and*

$$\|fg\|_1 \leqslant \|f\|_p \|g\|_q. \tag{6.4}$$

If $p = 1$, *equality iff* $|g| = \|g\|_\infty$, *a.e. on the set where* $f \neq 0$.

If $p = \infty$, equality iff $|f| = \|f\|_\infty$, a.e. on the set where $g \neq 0$.

If $1 < p < \infty$, equality iff there are nonnegative constants α, β such that $(\alpha, \beta) \neq (0, 0)$, $\alpha |f|^p = \beta |g|^q$, a.e.

Proof. **1**. The inequality easily follows if $p = 1$ or $p = \infty$. To see the equaliy, suppose $p = 1$, then $q = \infty$. (6.4) is equivalent to

$$\int |f| (\|g\|_\infty - |g|) \geqslant 0.$$

It is equality iff $|g| = \|g\|_\infty$, *a.e.* on the set where $f \neq 0$.

2. Suppose $1 < p$, $q < \infty$. The conclusion is obvious if $\|f\|_p = 0$ or $\|g\|_q = 0$. Hence we assume that $0 < \|f\|_p$, $\|g\|_q < \infty$. Using Young's inequality with

$$a = \left(\frac{|f|}{\|f\|_p} \right)^p, \; b = \left(\frac{|g|}{\|g\|_q} \right)^q,$$

we have

$$\frac{|fg|}{\|f\|_p\|g\|_q} \leqslant \frac{1}{p}\left(\frac{|f|}{\|f\|_p}\right)^p + \frac{1}{q}\left(\frac{|g|}{\|g\|_q}\right)^q, \ a.e.$$

Integrating on both sides gives

$$\int \frac{|fg|}{\|f\|_p\|g\|_q} d\mu \leqslant \frac{1}{p} + \frac{1}{q} = 1,$$

which is the desired inequality. The equality holds iff $a = b$, $a.e.$ i.e.,

$$\|g\|_q^q |f|^p = \|f\|_p^p |g|^q, \ a.e.$$

$\square$

A familiar case of Hölder inequality is the following.

Thm 35 (Cauchy–Schwarz inequality)**.** *f, $g \in L_2$, then $fg \in L_1$ and*

$$\|fg\|_1 \leqslant \|f\|_2\|g\|_2.$$

Thm 36 (Minkowski inequality). *$1 \leqslant p \leqslant \infty$, f, $g \in L_p$, then $f+g \in L_p$ and*

$$\|f+g\|_p \leqslant \|f\|_p + \|g\|_p. \tag{6.5}$$

If $p=1$ or $p=\infty$, equality iff $fg \geqslant 0$, a.e..

If $1 < p < \infty$, equality iff there are nonnegative constants α, β such that $(\alpha, \beta) \neq (0,0)$, $\alpha f = \beta g$, a.e.

Proof. **1**. The case $p=1$ or $p=\infty$ is immediate.

2. Suppose $1 < p < \infty$. Let $q > 1$, $\dfrac{1}{p} + \dfrac{1}{q} = 1$. By Hölder inequality

$$\begin{aligned}\|f+g\|_p^p = \int |f+g||f+g|^{p-1} &\leqslant_{(e1)} \int |f||f+g|^{p-1} + \int |g||f+g|^{p-1} \\ &\leqslant_{(e2)} \|f\|_p \big\||f+g|^{p-1}\big\|_q + \|g\|_p \big\||f+g|^{p-1}\big\|_q \\ &= \|f\|_p \|f+g\|_p^{p-1} + \|g\|_p \|f+g\|_p^{p-1}\end{aligned}$$

Here

$$\left\||f+g|^{p-1}\right\|_q = \left(\int \left(|f+g|^{p-1}\right)^q\right)^{1/q} = \left(\int |f+g|^p\right)^{1/q}$$
$$= \|f+g\|_p^{p/q} = \|f+g\|_p^{p-1}.$$

$(e1)$ is equality iff $fg \geqslant 0$, *a.e.*, $(e2)$ is equality iff there are nonegative constants a, b, c, d such that $(a,b) \neq (0,0)$, $(c,d) \neq (0,0)$,

$$a|f|^p = b\left(|f+g|^{p-1}\right)^q,\ c|g|^p = d\left(|f+g|^{p-1}\right)^q,\ a.e.$$

Hence

$$a|f| = b|f+g|,\ c|g| = d|f+g|,\ a.e.$$

The conclusion follows by combining the equality conditions of $(e1)(e2)$.

$\square$

Def 26. $0 < p < 1$, *let*

$$L_p \triangleq \left\{ f \in \mathscr{F} : \int |f|^p d\mu < \infty \right\}$$

and

$$\|f\|_p = \int |f|^p d\mu.$$

LEMMA 37. *Let* a, $b \in \mathbb{R}$, $0 < p < 1$. $|a+b|^p \leqslant |a|^p + |b|^p$.

PROOF. Since $||a|+|b||^p \leqslant |a|^p + |b|^p$ implies the desired inequality, we assume w.l.g. that a, b are of the same sign. Suppose $a \neq 0$, otherwise there is nothing to prove. Finally it suffices to show that

$$(1+s)^p \leqslant 1 + s^p, \ \ s \geqslant 0,$$

which is verified by elementary calculus. $\square$

Lemma 32 and Lemma 37 can be merged into the compact form,

$$|a+b|^p \leqslant C_p(|a|^p + |b|^p), \ 0 < p < \infty, \tag{6.6}$$

where $C_p = 2^{p-1} \vee 1$.

Thm 38. $0 < p < 1$, $\|f+g\|_p \leqslant \|f\|_p + \|g\|_p$.

6.2. Completeness.

Thm 39. *Let $0 < p \leqslant \infty$, L_p is complete.*

Proof for $p = \infty$. Let $f_n \in L_\infty$. Suppose that f_n is Cauchy. Given $k \geqslant 1$, there is n_k such that

$$\|f_m - f_n\|_\infty \leqslant \frac{1}{k},\ \forall m, n > n_k.$$

Hence there is a null set[1] A_k such that

$$|f_m - f_n| \leqslant \frac{1}{k},\ \forall x \in A_k^c,\ m, n > n_k.$$

Then $A = \bigcup_{k=1}^{\infty} A_k$ is a null set and $f_n(x)$ is Cauchy for each $x \in A^c$. Hence there exist f, $f_n \to f$ for $x \in A^c$. Let $m \to \infty$ in the above inequality we get

$$|f_n - f| \leqslant \frac{1}{k},\ \forall x \in A^c,\ n > n_k.$$

[1] A null set is a measurable set with measure zero.

So $f \in L_\infty$ and

$$\|f_n - f\|_\infty \leqslant \frac{1}{k}, \ \forall n > n_k.$$

Therefore f_n converges to f in L_∞. $\square$

Proof for $0 < p < \infty$. Let $f_n \in L_p$. Suppose that f_n is Cauchy in L_p,

$$\lim_{m,n\to\infty} \|f_m - f_n\|_p = 0. \tag{6.7}$$

We intend to show that $\lim_{n\to\infty} \|f_n - f\|_p = 0$ for some $f \in L_p$. Owing to (6.7), we have a subsequence $n_k \to \infty$ so that

$$\left\|f_{n_{k+1}} - f_{n_k}\right\|_p < \frac{1}{2^k}. \tag{6.8}$$

We claim that

(a) there is $h \in L_p$ such that $|f_{n_k}| \leqslant h$, $a.e.$
(b) $\lim_k f_{n_k} \to f$, $a.e.$ for some $f \in L_p$.
(c) $\lim_k \|f_{n_k} - f\|_p = 0$.

The conclusion of the Theorem clearly follows once (c) is proved, since a Cauchy sequence converges iff it has a convergent subsequence. Let

$$g_k = \sum_{i=1}^{k} \left| f_{n_{i+1}} - f_{n_i} \right|, \; g = \sum_{i=1}^{\infty} \left| f_{n_{i+1}} - f_{n_i} \right|.$$

Then $0 \leqslant g_k \uparrow g$ and $\|g_k\|_p \leqslant 1$ by (6.8) (Theorem 36 or Theorem 38). Using monontone convergence theorem,

$$\int g^p d\mu = \lim_k \int (g_k)^p d\mu \leqslant 1.$$

This shows $g \in L_p$ and that $g < \infty$, *a.e.* Therefore

$$f_{n_k} = f_{n_1} + \sum_{i=1}^{k} \left(f_{n_{i+1}} - f_{n_i} \right)$$

converges almost everywhere to some measurable function f and

$$|f_{n_k}| \leqslant |f_{n_1}| + g.$$

Let $k \to \infty$, we have

$$|f| \leqslant |f_{n_1}| + g, \ a.e.$$

hence $f \in L_p$. $(a)(b)$ follows with $h = |f_{n_1}| + g$. By inequality (6.6),

$$\begin{aligned} |f_{n_k} - f|^p &\leqslant C_p(|f_{n_k}|^p + |f|^p) \leqslant C_p(||f_{n_1}| + g|^p + |f|^p) \\ &\leqslant C_p(C_p(|f_{n_1}|^p + g^p) + |f|^p). \end{aligned}$$

Therefore (c) is a result of the dominated convergence theorem. $\square$

Corollary 1. (1) $0 < p < 1$, L_p *is a complete metric space.*
(2) $1 \leqslant p \leqslant \infty$, L_p *is a Banach space.*

6.3. L_p and weak convergence.

Thm 40. *Let* $0 < p < \infty$, $f_n \in L_p$, $f \in L_p$.
(1) $f_n \xrightarrow{L_p} f \implies f_n \xrightarrow{\mu} f$ *and* $\|f_n\|_p \to \|f\|_p$.
(2) $f_n \xrightarrow{a.e.} f$ *or* $f_n \xrightarrow{\mu} f$, *then*

$$\|f_n\|_p \to \|f\|_p \iff f_n \xrightarrow{L_p} f.$$

PROOF. **1**. To prove (1), use Markov inequality

$$\mu(|f_n - f| > \varepsilon) \leqslant \frac{1}{\varepsilon^p}\|f_n - f\|_p^p.$$

and the triangle inequality

$$\Big| \|f_n\|_p - \|f\|_p \Big| \leqslant \|f_n - f\|_p.$$

2. " $\Longleftarrow$ " of (2) is included in step **1**.

3. " $\Longrightarrow$ " of (2). In view of Theorem 20, it is enough to prove the case where $f_n \xrightarrow{a.e.} f$. Define

$$g_n = C_p(|f_n|^p + |f|^p) - |f_n - f|^p,$$

where $C_p = 2^{p-1} \vee 1$. Then $g_n \geqslant 0$ by inequality (6.6) and $\lim\limits_n g_n = 2C_p|f|^p$, *a.e.* Using Fatou's lemma

$$\begin{aligned}\int 2C_p|f|^p d\mu = \int \lim_n g_n d\mu &\leqslant \liminf_n \int g_n d\mu \\ &= \int 2C_p|f|^p d\mu - \limsup_n \int |f_n - f|^p.\end{aligned}$$

Canceling $\displaystyle\int 2C_p|f|^p d\mu$ from both side gives

$$\lim_n \int |f_n - f|^p = 0.$$

$\square$

Def 27. *$(E, \mathscr{F}, \mu)$ is a measure space. $1 \leqslant p < \infty$, $\dfrac{1}{p} + \dfrac{1}{q} = 1$, f_n converges weakly to f in L_p, denoted by $f_n \xrightarrow{w\text{-}L_p} f$, if*

$$\lim_n \int f_n g d\mu = \int f g d\mu, \ \forall g \in L_q.$$

μ is additionally assumed to be σ-finite if $p = 1$.

Thm 41. $1 \leqslant p < \infty$. *$f_n \xrightarrow{L_p} f$ implies $f_n \xrightarrow{w\text{-}L_p} f$.*

PROOF. By Hölder inequality (Theorem 34), $\forall g \in L_q$, q conjugate to p,

$$\int |f_n - f||g| d\mu \leqslant \|f_n - f\|_p \|g\|_q.$$

□

Thm 42. *$(E, \mathscr{F}, \mu)$ is a measure space. Let $1 < p < \infty$, $\{f_n\}$ bounded in L_p. If $f_n \xrightarrow{a.e.} f$ or $f_n \xrightarrow{\mu} f$ for some measurable f, then $f \in L_p$ and $f_n \xrightarrow{w\text{-}L_p} f$.*

PROOF. Let $g \in L_q$, q conjugate to p. As before, it is enough to prove it for $f_n \xrightarrow{a.e.} f$.

1. $f \in L_p$ is a consequence of Fatou's lemma,

$$\int |f|^p d\mu = \int \lim_n |f_n|^p d\mu \leqslant \liminf_n \int |f_n|^p d\mu \leqslant \sup_n \|f_n\|_{L_p} < \infty.$$

It follows that $\{f_n - f\}$ is bounded in L_p.

2. Fix $\varepsilon > 0$, let $\delta > 0$, define $A_\delta = \{x \in E : \delta \leqslant |g|^q \leqslant 1/\delta\}$ and write

$$\int |f_n - f||g| d\mu = \int_{A_\delta \cap B} + \int_{A_\delta \cap B^c} + \int_{A_\delta^c} .$$

Choose δ small so that

$$\int_{A_\delta^c} \leqslant \|f_n - f\|_p \big\| g 1_{A_\delta^c} \big\|_q < \frac{\varepsilon}{3}.$$

With δ fixed, we have

$$\int_{A_\delta \cap B^c} \leqslant \|f_n - f\|_p \|g 1_{A_\delta \cap B^c}\|_q < \frac{\varepsilon}{3},$$

as soon as $B \subset A_\delta$ is such that $\mu(A_\delta \cap B^c)$ is smaller than some ε'.

Note $|g| \leqslant 1/\delta^{1/q}$ on A_δ. Since $\mu(A_\delta)$ is finite by Markov inequality, so a subset $B \subset A_\delta$ can be chosen so that $\mu(A_\delta \cap B^c) < \varepsilon'$ and $|f_n - f|$

converges uniformly to 0 on $A_\delta \cap B$ (Theorem 19). Hence for large n,

$$\int_{A_\delta \cap B} \leqslant \frac{1}{\delta^{1/q}} \int_{A_\delta \cap B} |f_n - f| d\mu < \frac{\varepsilon}{3}.$$

$\square$

Note the above proof does not get through if $p = 1$ (so that $q = \infty$). The example below demonstates, in general, Theorem 42 does not for $p = 1$.

E.g. 12. *$E = (0,1)$ with the usual Lebesgue measure, $f_n = n1_{(0,1/n)}$. Clearly $\|f_n\|_1 = 1$, $f_n \xrightarrow{\mu} f = 0$. But with $g = 1 \in L_\infty$, $\lim_n \int f_n g d\mu = 1 \neq 0 = \int f g d\mu$, hence $f_n \xrightarrow{w\text{-}L_1} f$ does not hold.*

However we have

Thm 43. *$(E, \mathscr{F}, \mu)$ is a measure space. Let $\{f_n\} \in L_1$. Suppose $f_n \xrightarrow{a.e.} f$ or $f_n \xrightarrow{\mu} f$. Then*

$$f \in L_1,\ \|f_n\|_1 \to \|f\|_1 \iff f_n \xrightarrow{L_1} f.$$

Either of them gives $\int_A f_n d\mu \to \int_A f d\mu$, $\forall A \in \mathscr{F}$.

PROOF. The first conclusion is contained in Theorem 40. So $f_n \xrightarrow{w\text{-}L_2} f$ by Theorem 41. To complete the proof, take $1_A \in L_\infty$ as test function. $\square$

6.4. Uniform integrability. Let $(E, \mathscr{F}, \mu)$ be a measure space.

Def 28. $\mathcal{H} = \{f_t : t \in T\}$ *is uniformly integrable if*

$$\lim_{a\to\infty} \sup_{f\in\mathcal{H}} \int_{\{|f|\geqslant a\}} |f| d\mu = 0. \tag{6.9}$$

Def 29. $\mathcal{H} = \{f_t : t \in T\}$ *is absolutely continuous if,* $\forall \varepsilon > 0$, *there is* $\delta > 0$ *so that*

$$\sup_{f\in\mathcal{H}} \int 1_A |f| d\mu < \varepsilon \textit{ for any } A \textit{ with } \mu(A) < \delta.$$

Thm 44. *Suppose $(E, \mathscr{F}, \mu)$ is a measure space with μ finite. $\mathcal{H} = \{f_t : t \in T\}$ is uniformly integrable if and only if $\mathcal{H}$ is absolutely continuous and bounded in L_1.*

PROOF. **1**. If $\mathcal{H}$ is uniformly integrable, $\forall \varepsilon > 0$, there is $a_0 > 0$ so that
$$\sup_{f \in \mathcal{H}} \int_{\{|f| \geqslant a\}} |f| d\mu \leqslant \frac{\varepsilon}{2}, \ \forall a \geqslant a_0.$$
For any measurable A, $a \geqslant a_0$,
$$\begin{aligned}\sup_{f \in \mathcal{H}} \int 1_A |f| d\mu &\leqslant \sup_{f \in \mathcal{H}} \int_{\{|f| < a\}} 1_A |f| d\mu + \sup_{f \in \mathcal{H}} \int_{\{|f| \geqslant a\}} 1_A |f| d\mu \\ &\leqslant a\mu(A) + \sup_{f \in \mathcal{H}} \int_{\{|f| \geqslant a\}} |f| d\mu \leqslant a\mu(A) + \frac{\varepsilon}{2}.\end{aligned}$$
That $\mathcal{H}$ is bounded in L_1 follows by setting $A = E$ and using the fact that μ is finite. Fix $a \geqslant a_0$. For any A with $\mu(A) \leqslant \varepsilon/(2a)$, we get that $\sup_{f \in \mathcal{H}} \int 1_A |f| d\mu$ is bounded from above by ε, hence the absolute continuity.

2. Suppose that $\mathcal{H}$ is absolutely continuous and bounded in L_1. Denote the uniform L_1 bound of $\mathcal{H}$ by M. By Markov inequality, $\forall a > 0$,

$$\mu(|f| > a) \leqslant \frac{1}{a}\int |f| d\mu \leqslant \frac{1}{a}M, \ \forall f \in \mathcal{H}.$$

$\forall \varepsilon > 0$, by absolute continuity, $\sup_{f\in\mathcal{H}} \int 1_A |f| d\mu < \varepsilon$ as soon as $\mu(A)$ is less than some $\delta > 0$. Fix a with $M/a < \delta$. Then setting $A = \mu(|f| > a)$ gives the uniform integrability. $\square$

Thm 45 (Vitali convergence theorem)**.** *Suppose that μ is finite, $f_n \xrightarrow{a.e.} f$ or $f_n \xrightarrow{\mu} f$.*

(1) *If $\{f_n\}$ is uniformly integrable, then $f \in L_1$ and*

$$\int f_n d\mu \to \int f d\mu. \tag{6.10}$$

(2) *f_n, f are nonnegative integrable, then (6.10) implies that $\{f_n\}$ is uniformly integrable.*

Proof. The proof is given for $f_n \xrightarrow{a.e.} f$.

1. If f_n is uniformly integrable, then f is integrable by Theorem 44 and Fatou's lemma. Define

$$f_{n,a} = 1_{\{|f_n|<a\}} f_n,\ \ f_a = 1_{\{|f|<a\}} f.$$

It follows that $f_{n,a} \to f_a$, $a.e.$ provided $\mu(|f| = a) = 0$. By bounded dominated convergence,

$$\int f_{n,a} d\mu \to \int f_a d\mu.$$

Writing

$$\int_{\{|f_n|\geqslant a\}} f_n d\mu = \int f_n d\mu - \int f_{n,a} d\mu \tag{6.11}$$

and

$$\int_{\{|f|\geqslant a\}} f d\mu = \int f d\mu - \int f_a d\mu, \tag{6.12}$$

we see that

$$
\begin{aligned}
&\limsup_n \left| \int f_n d\mu - \int f d\mu \right| \\
&\leqslant \limsup_n \left| \int f_{n,a} d\mu - \int f_a d\mu \right| + \sup_n \int_{\{|f_n|\geqslant a\}} |f_n| d\mu + \int_{\{|f|\geqslant a\}} |f| d\mu \\
&= \sup_n \int_{\{|f_n|\geqslant a\}} |f_n| d\mu + \int_{\{|f|\geqslant a\}} |f| d\mu.
\end{aligned}
$$

Note $\mu(|f| = a) = 0$ for all but countably many a. Sending $a \to \infty$ proves (6.10).

2. Suppose f_n, f are nonnegative integrable and (6.10) holds. Write

$$
\int_{\{|f_n|\geqslant a\}} f_n d\mu = \int_{\{|f|\geqslant a\}} f d\mu + \left(\int_{\{|f_n|\geqslant a\}} f_n d\mu - \int_{\{|f|\geqslant a\}} f d\mu \right).
$$

Since f is integrable, the first term is less than $\varepsilon/2$ when a is larger than some a_0. If $\mu(|f| = a) = 0$, (6.11) and (6.12) indicate the term in the bracket is also less than $\varepsilon/2$ when n is larger than some n_0.

Therefore,

$$\sup_{n>n_0}\int_{\{|f_n|\geqslant a\}} f_n d\mu \leqslant \varepsilon,\ \forall a > a_0 \text{ with } \mu(|f| = a) = 0.$$

Since the finite family $\{f_1, ..., f_{n_0}\}$ is uniformly integrable, the uniform integrability of $\{f_n, n \geqslant 1\}$ follows. $\square$

Additional details on the proof of Theorem 45. Suppose $|f_n(x)| \to |f(x)| < a$. Then for large n, $|f_n(x)| < a$. So $1_{\{|f_n|<a\}}$ and $1_{\{|f|<a\}}$ are both equal to 1, it follows $f_{n,a} \to f_a$ at x. The same is true for x with $|f(x)| > a$. If $|f(x)| = a \neq 0$, then $f_{n,a}(x) \to f_a(x)$ may not happen, since in this case $f_a(x) = 0$ while there could be a subsequence n_k with $f_{n_k}(x) < a$ so that

$$f_{n_k,a}(x) = f_{n_k}(x) \to f(x) \neq 0.$$

But if the set $\{x : |f(x)| = a\}$ has zero μ-measure, then $f_{n,a} \to f_a$, $a.e.$ Fortunately the set of a for which $\mu(|f| = a)$ is not zero is at most countable. Indeed, let

$$F(x) = \mu(|f| \leqslant x).$$

Then $F(x)$ is non-decreasing, hence has at most countably many discontinuities. F is discontinuous at $x = a$ if and only if

$$\mu(|f| = a) = F(a) - F(a-) \neq 0.$$

This verifies that $\mu(|f| = a) = 0$ for all but countably many a.

COROLLARY 2. *Suppose that μ is finite, f_n, f are integrable. If $f_n \xrightarrow{a.e.} f$ or $f_n \xrightarrow{\mu} f$, then these are equivalent:*
(1) *$\{f_n\}$ is uniformly integrable;*
(2) $\displaystyle\int |f_n - f| d\mu \to 0$;
(3) $\displaystyle\int |f_n| d\mu \to \int |f| d\mu$.

6.5. Summary of various convergences.

$$
\begin{array}{ccccccc}
f_n \xrightarrow{\mu} f & \overset{\text{sub subseq}}{\Longleftrightarrow} & f_n \xrightarrow{a.u} f & & \\
 & \text{Thm } 20 & & & \\
\text{Markov} \Uparrow & & & & \\
f_n \xrightarrow{L_p} f & \begin{array}{c}\overset{\text{has a subseq}}{\Longrightarrow} \\ \underset{\|f_n\|_p \to \|f\|_p}{\Longleftarrow}\end{array} & f_n \xrightarrow{a.e.} f & \begin{array}{c}\overset{\mu \text{ finite}}{\Longleftrightarrow} \\ \Longleftarrow\end{array} & f_n \xrightarrow{a.u} f \\
 & \text{Thm } 40 & & \text{Thm } 19 & \Downarrow \text{Thm } 18 \\
 & & f_n \xrightarrow{d} f & \overset{\text{in a prob space}}{\Longleftarrow} & f_n \xrightarrow{\mu} f \\
 & & & \text{Thm } 24 &
\end{array}
$$

7. 概率空间的积分

7.1. Expected value. $(\Omega, \mathscr{F}, P)$ is a probability space, X a r.v.

Def 30. *Expectation, written* EX,

$$EX = \int X dP.$$

Suppose X is discrete, i.e., X takes values in a finite or infinitely countable *distinct* sequence $\{x_1, x_2, ...\}$. Then its expectation ($\int X dP$ computed according to (5.1)) equals

$$EX = \sum_i x_i P(X = x_i).$$

The mapping $i \mapsto P(X = x_i)$ is called the probability mass function of X. If $Y = g(X)$ for some measurable function g, then Y is discrete with values in, say, $\{y_1, y_2, ...\}$. The expectation of Y, computed in the same way as EX, is

$$EY = \sum_i y_i P(Y = y_i).$$

To calculate EY, we first need to find its probability mass function $i \mapsto P(Y = y_i)$. This can be complicated, and it is avoided by using the "*law of the unconscious statistician*",

$$EY = \sum_i g(x_i) P(X = x_i).$$

This turns out to be a change of variables formula (see also Theorem 48).

Thm 46 (Change of variables formula)**.** *Let* $(\Omega, \mathscr{F}, P)$ *be a probability space,* X *a r.v, and* $g \in \mathscr{B}_{\mathbb{R}}$*. If* $g \geqslant 0$ *or* $\int_\Omega |g(X)| dP < \infty$, *then*

$$Eg(X) = \int_\Omega g(X) dP = \int_{\mathbb{R}} g(x) d\mu_X. \tag{7.1}$$

Here $\mu_X(A) = PX^{-1}(A) = P(X \in A)$, $\forall A \in \mathscr{B}_{\mathbb{R}}$ is the probability induced by X (section 5.3).

PROOF. **1**. The nonnegative case $g \geqslant 0$. If $g = 1_A$, then $g(X(\omega)) = 1_A(X(\omega)) = 1_{X^{-1}(A)}(\omega)$, so (7.1) reduces to the definition of μ_X. By linearity, (7.1) holds for simple functions. If g_n are simple functions such that $0 \leqslant g_n(x) \uparrow g(x)$, then $0 \leqslant g_n(X(\omega)) \uparrow g(X(\omega))$, then (7.1) follows by monontone convergence theorem.

2. The case $\displaystyle\int_\Omega |g(X)|dP < \infty$. Applying step **1** to $|g(X)|$ shows that g is integrable with respect to μ_X, hence the integrability of g^+, g^-, and (7.1) follows from subtracting $Eg^-(X) = \displaystyle\int_{\mathbb{R}} g^-(x)d\mu_X$ from $Eg^+(X) = \displaystyle\int_{\mathbb{R}} g^+(x)d\mu_X$. $\square$

The probability μ_X equals (as a result of the uniqueness Theorem 11) the measure μ constructed from the distribution function F of X : $\mu((a,b]) = F(b) - F(a)$, $\forall a, b$. The measure μ is called a

Lebesgue-Stieltjes measure and its integral is the Lebesgue-Stieltjes integral (section 7.3). The above formula thus relates integral on a probability space to Lebesgue-Stieltjes integral over $\mathbb{R}$. The rightmost term of (7.1) is also written as $\int g dF$, i.e.

$$Eg(X) = \int_{\mathbb{R}} g(x) dF.$$

REMARK 2. *An implication of Theorem 46 is that the integration (e.g. the expection and variance) of a random variable is a distributional property, i.e., it depends on the random variable only through its distribution. This lays the basis for applying probability theory tools such as Skorohod Theorem (Theorem 26).*

Def 31. *Variance, written* $Var(X)$,

$$Var(X) = \int (X - EX)^2 dP = E\big[(X - EX)^2\big].$$

It is easy to see that

$$Var(X) = EX^2 - (EX)^2.$$

Def 32. *k-th moment, $k = 1, 2, ...$,*

$$E(X^k) = \int X^k dP.$$

E.g. 13 (Bernoulli distribution)**.** *Let $0 < p < 1$. $X \sim Bernoulli(p)$ if $P(X = 1) = p$, $P(X = 0) = 1 - p$. Then*

$$EX = 1 \cdot p + 0 \cdot (1 - p) = p.$$

$$Var(X) = EX^2 - (EX)^2 = p - p^2 = p(1 - p).$$

E.g. 14 (Poisson distribution)**.** *Let $\lambda > 0$. $X \sim Poisson(\lambda)$ if*

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \ k = 0, 1, 2, ...$$

Then

$$EX = \sum_{k=0}^{\infty} k \cdot e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} = \lambda.$$

$$E(X(X-1))=\sum_{k=0}^{\infty}k(k-1)\cdot e^{-\lambda}\frac{\lambda^k}{k!}=e^{-\lambda}\sum_{k=2}^{\infty}\frac{\lambda^k}{(k-2)!}=\lambda^2.$$

Hence $EX^2=\lambda^2+\lambda$, *and*

$$Var(X)=\left(\lambda^2+\lambda\right)-\lambda^2=\lambda.$$

7.2. Properties of expectation. X, Y are random variables. The following are immediate from section 6.1.

Jensen inequality: if X integrable, φ convex, then

$$\varphi(EX)\leqslant E\varphi(X).$$

Hölder inequality: if $p,q\geqslant 1, 1/p+1/q=1$, then

$$E|XY|\leqslant \|X\|_p\|Y\|_q.$$

Minkowski inequality: if $p\geqslant 1$, then

$$\|X+Y\|_p\leqslant \|X\|_p+\|Y\|_p.$$

Thm 47. $0<s<t<\infty$, X *is a r.v. Then* $\|X\|_s\leqslant\|X\|_t$.

Proof. By Hölder inequality with $p = \frac{t}{s}$, $q = \frac{t}{t-s}$,

$$\|X\|_s^s = E|X|^s \leqslant (E|X|^{sp})^{1/p}(E1^q)^{1/q} = \left(E|X|^t\right)^{s/t} = \|X\|_t^s.$$

$\square$

E.g. 15. *If X has $EX^2 < \infty$, then its expectation and variance exist, since $E|X| \leqslant \|X\|_2 < \infty$, and*

$$0 \leqslant Var(X) \leqslant EX^2.$$

7.3. Lebesgue-Stieltjes and Riemann-Stieltjes integrals. Let G be a **generalized distribution function**, i.e., nondecreasing, right-continuous on $\mathbb{R}$. There is a unique measure μ such that

$$\mu((a,b]) = G(b) - G(a), \ \forall a, b. \tag{7.2}$$

The measure μ constructed this way is called a **Lebesgue-Stieltjes measure**. Integration with respect to Lebesgue-Stieltjes measure is called **Lebesgue-Stieltjes integral**, denoted by $\int f d\mu$ or $\int f dG$.

Remark 3. *Under suitable conditions (see below),* $\int f dG$ *may be interpreted as Riemann-Stieltjes integral. Since this does not provide anything new in the context of general measure theory,* $\int f dG$ *is best understood as a notional variant of* $\int f d\mu$.

Here we recall a few facts about Riemann-Stieltjes integration. Let G be the function as in (7.2), f a bounded function on $[a,b]$. Corresponding to each partition $\mathcal{P} : a = x_0 < x_1 < \cdots < x_n = b$, we consider

$$L(\mathcal{P}, f) = \sum_{i=1}^{n} \inf_{x \in [x_{i-1}, x_i]} f(x) \Delta G_i, \; U(\mathcal{P}, f) = \sum_{i=1}^{n} \sup_{x \in [x_{i-1}, x_i]} f(x) \Delta G_i.$$

Here $\Delta G_i = G(x_i) - G(x_{i-1})$. Define

$$R_* f = \sup_{\mathcal{P}} L(\mathcal{P}, f), \; R^* f = \inf_{\mathcal{P}} U(\mathcal{P}, f).$$

If $R_*f = R^*f$, then f is Riemann-Stieltjes integrable with respect to G, the common value, written $(R\text{-}S)\int f$, is called the Riemann-Stieltjes integral. For simplicity we have omitted the dependence of the integral on G in the notations.

A sufficient condition for Riemann-Stieltjes integrability is this: Suppose f is bounded on $[a,b]$, has at most finitely many discontinuities, G is continuous at every point where f is discontinuous. Then f is Riemann-Stieltjes integrable with respect to G.

E.g. 16. *If $a < s < b$, f is bounded on $[a,b]$, continuous at s and $G(x) = 1_{[s,\infty)}(x)$. Then*

$$(R\text{-}S)\int_a^b f dG = f(s).$$

Indeed, consider paritions $\mathcal{P} = \{x_0, x_1, x_2, x_3\}$, $a = x_0$ and $x_1 < x_2 = s < x_3 = b$. Then $\Delta G_2 = 1$, $\Delta G_i = 0$ if $i \neq 2$,

$$L(\mathcal{P}, f) = \inf_{x\in[x_1,x_2]} f(x),\ U(\mathcal{P}, f) = \sup_{x\in[x_1,x_2]} f(x).$$

Since f is continuous at s, we see that $L(\mathcal{P}, f)$ and $U(\mathcal{P}, f)$ converge to $f(s)$ as $x_1 \to s$.

Thm 48. *Suppose $c_n \geqslant 0$, $\sum c_n < \infty$, $\{s_n\}$ is a sequence of distinct points in (a, b), and*

$$G(x) = \sum_{n=1}^{\infty} c_n 1_{[s_n,\infty)}(x).$$

If f is continuous on $[a, b]$, then

$$(R\text{-}S) \int_a^b f dG = \sum_{n=1}^{\infty} c_n f(s_n).$$

PROOF. Exercise. $\square$

If we denote by $L_* f$ the integral in (5.1) with the G-induced Lebesgue-Stieltjes measure in the role of μ, and by $L^* f$ the integral in (5.2). Then

$$R_* f \leqslant L_* f \leqslant L^* f \leqslant R^* f.$$

Therefore if, for instance, f is continuous on $[a, b]$, then it is Riemann-Stieltjes integrable, hence Lebesgue-Stieltjes integrable.

7.4. L_p convergence and uniform integrability.

Thm 49. *$(\Omega, \mathscr{F}, P)$ is a probability space, $0 < p < \infty$, $X_n \in L_p$, $X \in \mathscr{F}$. If $X_n \xrightarrow{P} X$, then these are equivalent:*
(1) *$\{|X_n|^p\}$ is uniformly integrable;*
(2) *$X \in L_p$, $E(|X_n - X|^p) \to 0$;*
(3) *$X \in L_p$, $E(|X_n|^p) \to E(|X|^p)$.*

Proof. **1**. Observe that $X \in L_p$ by Theorem 45, hence $\{|X_n - X|^p\}$ is uniformly integrable since $|X_n - X|^p \leqslant C_p(|X_n|^p + |X|^p)$ where $C_p = 2^{p-1} \vee 1$. Note also that $|X_n - X|^p \xrightarrow{P} 0$. Therefore (1) implies (2) is a consequence of Theorem 45 with $f_n = |X_n - X|^p$.

2. (2) implies (3) because $\left| \|X_n\|_p - \|X\|_p \right| \leqslant \|X_n - X\|_p$, $0 < p < \infty$ (Theorem 36, Theorem 38).

3. (3) implies (2) follows from an application of Theorem 45 with $f_n = |X_n|^p$. □

We notice another criterion for uniform integrability, in addition to Theorem 44.

LEMMA 50. *Let* $(\Omega, \mathscr{F}, P)$ *be a probability space,*

$$\mathcal{H} = \{X_t : t \in T,\ E|X_t| < \infty\}.$$

Suppose that $g \geqslant 0$ *is an increasing function on* $[0, \infty)$ *such that*

$$\lim_{s\to\infty} \frac{g(s)}{s} = \infty$$

and

$$\sup_{X\in\mathcal{H}} \int g(|X|)dP < \infty.$$

Then $\mathcal{H}$ *is uniformly integrable.*

PROOF. $\forall \varepsilon > 0$. Fix $a > 0$ so that

$$\frac{1}{a} \sup_{X\in\mathcal{H}} \int g(|X|)dP < \varepsilon.$$

There is $s_0 > 0$ such that $g(s) \geqslant as$ for all $s \geqslant s_0$. Hence, $\forall X \in \mathcal{H}$, $s \geqslant s_0$,

$$\int_{\{|X| \geqslant s\}} |X| dP \leqslant \frac{1}{a} \int_{\{|X| \geqslant s\}} g(|X|) dP \leqslant \frac{1}{a} \sup_{X \in \mathcal{H}} \int g(|X|) dP < \varepsilon.$$

$\square$