思想：$\begin{cases} \text{Dykin } \pi\text{-}\lambda \\ \text{逼近} \end{cases}$

Reference：① 钟开莱
② Paul R. Halmos. Measure Theory (工具书)
③ Davis Williams: Probability with Martingale.

简介：

本科概率论 v.s. 高等概率

① 概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ —— Ω: 样本空间，$\mathcal{F}$: 事件域（σ-代数），$\mathbb{P}$: 概率测度（定义域为$\mathcal{F}$的函数）

1933. Kolmogorov 公理化体系.

1812. 古典概型 Laplace《分析概率论》

$(\Omega, \mathcal{F}, \mathbb{P})$ 测度空间

- Dykin π-λ thm
- Halmos 单调类 thm
- Carathéodory 延拓 thm

② r.v. $X(\omega)$

$F(x)=\mathbb{P}(X\le x)$

$F(x)$ is A.C. $\Rightarrow dF(x)=p(x)dx$

r.v. $X(\omega):(\Omega,\mathcal{F})\to(S,\mathcal{S})$ 是"可测"函数（$\mathcal{S}$: σ-代数）

$\boxed{\text{i.e. } \forall B\in\mathcal{S}, X^{-1}(B)\in\mathcal{F}}$，$X^{-1}(B)=\{\omega\in\Omega, X(\omega)\in B\}$

r.v. 分布 $P_X(B)\triangleq\mathbb{P}(X^{-1}(B)), \forall B\in\mathcal{S}$

$\Rightarrow (S,\mathcal{S},P)$ 概率空间.

若 $(S,\mathcal{S})=(\mathbb{R},\mathcal{B}_{\mathbb{R}})$，则 $F(x)=P_X((-\infty,x]), x\in\mathbb{R}$

r.v. X, Y

分布 P_X, P_Y （$P_X=P_Y$ 同分布）

$\Updownarrow$

分布函数 $F_X(x)$, $F_Y(x)$ （$F_X=F_Y$ 同分布）.

③ 数学期望 $EX=\int_{\mathbb{R}} xp(x)dx$

r.v. X 关于$\mathbb{P}$的积分 $EX=\int_\Omega X(\omega)\mathbb{P}(d\omega)$

$\overset{\text{换元}}{=}\int_S x P_X(dx)$

$\overset{S=\mathbb{R}}{=}\int_{\mathbb{R}} xF_X(dx)$

$F_X(dx)\ll dx \quad \overset{p(x)\triangleq\frac{F_X(dx)}{dx}}{=} \int_{\mathbb{R}} xp(x)dx$

条件数学期望 $E[X|Y]$, $E[X|Y=y]$

$E[X|\mathcal{G}]$, $\mathcal{G}\subset\mathcal{F}$ 为σ-代数

- 引：Radon-Nikodym thm
- 鞅论，回归分析，Filter理论.

④ r.v. 收敛性.

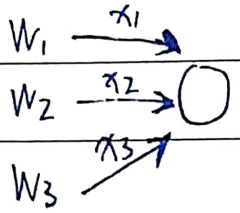
a.e.收，依P收，

依d收敛 $\Leftrightarrow$ 概率测度的弱收敛 $\Leftarrow$ Tightness

- 连续映射thm, Skorokhod 表示 thm, Vitali收敛thm, Helly thm, Prohov thm.

「A Mean-Field View of the Landscape of Two-Layers Neural-Networks」

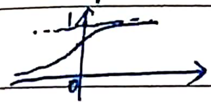
- Rosenblatt 感知器 (percetron)



$y=\mathbb{1}\{wx+\overset{\text{bias}}{b}>0\}$ 或 $y=\tilde{\sigma}(wx+b)$

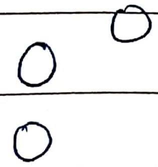
其中 $\tilde{\sigma}(z)=\mathbb{1}_{(0,+\infty)}(z)$

- Sigmoid 感知器: $\tilde{\sigma}(z)=\frac{1}{1+e^{-z}}, z\in\mathbb{R}$.



- 数据 $(x_i, y_i)\in\mathbb{R}^d\times\mathbb{R}$, $i\geq 1$. i.i.d $\sim P$.

  (x_i: Feature vector, y_i: Label)



- 参数 $\theta=(\theta_1,\cdots,\theta_N)$, N表示"O"个数, $\theta_i\in\mathbb{R}^m$.

  $\sigma(x,\theta_i)=\tilde{\sigma}(w_i x+b_i)$, $\theta_i=(w_i,b_i)$

- Output of 2-Layer N-N: $\hat{y}(x,\theta)=\frac{1}{N}\sum_{i=1}^{N}\sigma(x,\theta_i)$ (θ: 未知.)

- 损失函数 loss func $L:\mathbb{R}\times\mathbb{R}\to\mathbb{R}_+$

  广义误差 population risk: $R_N(\theta)\triangleq\mathbb{E}[L(y,\hat{y}(x,\theta))]$

Question: 找 θ^* $\Rightarrow$ $\inf_{\theta}R_N(\theta)=R_N(\theta^*)$

实际: 用SGD做:

$$\theta_i^{k+1}=\theta_i^k+2s_k(y_k-\hat{y}(x_k,\theta^k))\nabla_{\theta_i}\sigma(x_k,\theta_i^k)$$

- $R_N(\theta)\overset{L(y_1,y_2)=|y_1-y_2|^2}{=}\mathbb{E}\left|y-\frac{1}{N}\sum_{i=1}^{N}\sigma(x,\theta_i)\right|^2$

$$=\mathbb{E}y^2-\frac{2}{N}\sum_{i=1}^{N}\mathbb{E}[y\sigma(x,\theta_i)]+\frac{1}{N^2}\sum_{i,j=1}^{N}\mathbb{E}[\sigma(x,\theta_i)\sigma(x,\theta_j)]$$

Define $V(\theta_i)\triangleq-\mathbb{E}[y\sigma(x,\theta_i)]$, $U(\theta_i,\theta_j)=\mathbb{E}[\sigma(x,\theta_i)\sigma(x,\theta_j)]$, $R_\#\triangleq\mathbb{E}y^2$

则 $R_N(\theta)=R_\#+\frac{2}{N}\sum_i V(\theta_i)+\frac{1}{N^2}\sum_{i,j}U(\theta_i,\theta_j)$

Define $\rho^{(N)}\triangleq\frac{1}{N}\sum_{i=1}^{N}\delta_{\theta_i}$ ⇐ Dirac-Delta测度 (概率测度)

即 $\forall A\in\mathcal{B}_{\mathbb{R}^m}$, $\rho^{(N)}(A)=\frac{1}{N}\sum_i\delta_{\theta_i}(A)=\frac{1}{N}\sum_i\mathbb{1}_A(\theta_i)$.

则 $\rho^{(N)}$ 为 empirical probability measure.

于是, $\frac{1}{N}\sum_i V(\theta_i) = \int_{\mathbb{R}^m} V(\theta)\,\rho^{(N)}(d\theta)$

$\frac{1}{N^2}\sum_{i,j} U(\theta_i,\theta_j) = \int_{\mathbb{R}^m\times\mathbb{R}^m} U(\theta,\hat\theta)\,\rho^{(N)}(d\theta)\,\rho^{(N)}(d\hat\theta)$

$\Rightarrow R_N(\theta) = R_\# + 2\int_{\mathbb{R}^m} V(\theta)\,\rho^{(N)}(d\theta) + \int_{\mathbb{R}^m\times\mathbb{R}^m} U(\theta,\hat\theta)\,\rho^{(N)}(d\theta)\,\rho^{(N)}(d\hat\theta).$

Theorem. Under Assumptions, $\inf_\theta R_N(\theta) = \inf_\rho R(\rho) + O(\frac{1}{N})$, 其中

$R(\rho) = R_\# + 2\int V(\theta)\,\rho(d\theta) + \int (\theta,\hat\theta)\,\rho(d\theta)\,\rho(d\hat\theta)$

ρ是概率测度, $\rho \mapsto R(\rho)$ convex.

SGD θ_i^k 极限的解释: $\rho_k^{(N)} \triangleq \frac{1}{N}\sum_{i=1}^N \delta_{\theta_i^k}$, 取 step-size $s_k = \varepsilon\,\xi(k\varepsilon)$, $\varepsilon>0$, ξ光滑.

Theorem: Under Assumptions, $\rho^{(N)}_{[t/\varepsilon]} \underset{\varepsilon\to 0}{\overset{N\to\infty}{\Rightarrow}} \rho_t^{(33)}$, 其中

$\partial\rho_t = 2\xi(t)\nabla_\theta\left(\rho_t \nabla_\theta \Psi(\theta;\rho_t)\right)$, $\Psi(\theta;\rho) \triangleq V(\theta) + \int U(\theta,\tilde\theta)\,\rho(d\tilde\theta)$

Ex. $\forall \alpha \in I$, $\mathcal{F}_\alpha$ 为 σ-alg. 则 $\bigcap_{\alpha\in I} \mathcal{F}_\alpha$ 为 σ-alg.

Ex. $\mathcal{H} \subset \mathcal{G}$ 且 $\mathcal{H}$, $\mathcal{G}$ 均为 σ-alg. $\forall H \in \mathcal{H}$, 定义 $\mathcal{H}^H \triangleq \{A \in \mathcal{G} : A \cap H \in \mathcal{H}\}$.

证明. (a) $\mathcal{H}^H$ 为 σ-alg

(b) $H \to \mathcal{H}^H$ 单减

(c) $\forall H, H' \in \mathcal{H}$, $\mathcal{H}^{H \cup H'} = \mathcal{H}^H \cap \mathcal{H}^{H'}$.

第一章 概率空间与随机变量

§1.1 σ-代数. $(\Omega, \mathcal{F}, P)$.

设样本空间 Ω 为任一空间.

Def. (代数). 设 $\mathcal{A}$ 为 Ω 中某些子集所形成的集类. 若

(1) $\Omega \in \mathcal{A}$

(2) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$ (补封闭)

(3) $A_i \in \mathcal{A}, i=1,\cdots,n \Rightarrow \bigcup_{i=1}^{n} A_i \in \mathcal{A}$ (有限并封闭.)

则称 $\mathcal{A}$ 为一个代数.

RMK. $\mathcal{A} = \{\phi, \Omega\}$ 平凡代数

$\mathcal{A} = 2^\Omega$ 最大代数.

$\mathcal{A} = \{\phi, A, A^c, \Omega\}$ 为一个代数 $(A \subset \Omega)$

e.g. 设 $\Omega = (0,1]$

$\mathcal{A} = \{\phi, \bigcup_{i=1}^{n} (a_i, b_i], 0 \le a_1 < b_1 < \cdots < a_n < b_n \le 1, n \ge 1\}$. 为一个代数.

注意, 无穷并不封闭.

Def. (σ-代数) 设 $\mathcal{F}$ 为 Ω 中某些子集所形成的集类. 若

(1) $\Omega \in \mathcal{F}$

(2) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$

(3) $A_i \in \mathcal{F}, i \ge 1 \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$ (可列并封闭)

则称 $\mathcal{F}$ σ-代数.

RMK. $\mathcal{F} = \{\phi, \Omega\}$ 平凡代数.

$\mathcal{F} = 2^\Omega$ 最大 σ-代数 (事件域)

σ-代数 $\overset{\Rightarrow}{\not\Leftarrow}$ 代数.

· 过滤 (Filtration): 一列(族) 单增 σ-代数 $\{\mathcal{F}_n\}_{n=0,1,\cdots}$, $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots$

(信息流). $\{\mathcal{F}_t\}_{t\ge 0}$, $\mathcal{F}_s \subset \mathcal{F}_t$, $s<t$

e.g. (Borel.) $\Omega=\mathbb{R}$, $\mathcal{A}=\{\mathbb{R}中开集\}$

$\mathcal{B}_{\mathbb{R}} \triangleq \sigma(\mathcal{A})$ 为 Borel σ-代数.

称每个 Borel σ-代数中的集合为 Borel集. (包括: 开集, 闭集, G_δ, F_σ ...)

$G_\delta(F_\sigma)$: 可数个开(闭)集的交(并)

则 $\mathcal{B}_{\mathbb{R}} \subsetneq 2^{\mathbb{R}}$.

(拓展: Ω 为拓扑空间, $\mathcal{B}_\Omega=\sigma(\{\Omega中开集\ O\})$)

Pf of λ-类等价定义:

((ii)' + (iii)' $\Rightarrow$ (i) + (iii)):

1° 证(ii): 设 $A\subset B\in\mathcal{L}$, 则 $A\cup B^c \overset{(ii)',(iii)'}{\in} \mathcal{L}$, $B\setminus A=(A\cup B^c)^c \overset{(ii)'}{\in}\mathcal{L}$

2° 证(iii): 令 $B_i=A_i\setminus A_{i-1}$ (其中 $A_0=\phi$), 则 $\bigcup_{i=1}^{\infty}A_i=\bigcup_{i=1}^{\infty}B_i \overset{(iii)'}{\in}\mathcal{L}$

((ii) + (iii) $\Rightarrow$ (ii)' + (iii)'):

1° (ii)' 显然

2° (iii)': 设 $A_n\in\mathcal{L}$ 且 $A_i\cap A_j=\phi$, 则 $A_i\subset A_j^c \overset{(ii)}{\Rightarrow} A_j^c\setminus A_i\in\mathcal{L}$
$= A_j^c\cap A_i^c$
$\Rightarrow A_i\cup A_j\in\mathcal{L}$

令 $B_n=\bigcup_{i=1}^{n}A_i$, 则 $B_n\uparrow\in\mathcal{L}$ $\therefore \bigcup_{i=1}^{\infty}A_i=\bigcup_{i=1}^{\infty}B_i\overset{(iii)}{\in}\mathcal{L}$ #

Pf of Dykin π-λ thm:

outline: 设 $\lambda(\mathcal{A})$ 为由 $\mathcal{A}$ 生成的 λ-类, 则 $\lambda(\mathcal{A})\subset\mathcal{L}$.

若 $\lambda(\mathcal{A})$ 为 π-类 $\Rightarrow$ ($\lambda+\pi=\sigma$) $\sigma(\mathcal{A})\subset\lambda(\mathcal{A})$ $\Rightarrow \sigma(\mathcal{A})\subset\mathcal{L}$

只需证: $\lambda(\mathcal{A})$ 为 π-类.

令 $D_2\triangleq\{B\in\lambda(\mathcal{A}): B\cap C\in\lambda(\mathcal{A}), \forall C\in\lambda(\mathcal{A})\}$, 即证 $\lambda(A)\subset D_2$

只需证 D_2 为 λ类 (HW) 且 $\mathcal{A}\subset D_2$. $\Leftarrow \forall B\in\mathcal{A}, \forall C\in\lambda(\mathcal{A})$, 有 $B\cap C\in\lambda(\mathcal{A})$

令 $D_1\triangleq\{C\in\lambda(\mathcal{A}): B\cap C\in\lambda(\mathcal{A}), \forall B\in\mathcal{A}\}$, 则 D_1 为 λ类 (HW) 且 $\mathcal{A}\subset D_1 \Rightarrow \lambda(\mathcal{A}) = D_1$

只需 $\forall C\in\mathcal{A}, \forall B\in\mathcal{A}$, 有 $B\cap C\in\lambda(\mathcal{A})$

由 $\mathcal{A}$ 为 π类, $B\cap C\in\mathcal{A}\subset\lambda(\mathcal{A})$.

Def. (生成的 σ-代数) 设 $\mathcal{A}\subset 2^\Omega$, $\sigma(\mathcal{A})\triangleq\bigcap_{\substack{\mathcal{A}\subset\mathcal{G}为\sigma-代数\\ \mathcal{G}\subset 2^\Omega}}\mathcal{G}$ 称为由 $\mathcal{A}$ 生成的(最小) σ-代数

RMK. • $\sigma(\mathcal{A})$ 是包含 $\mathcal{A}$ 的最小 σ-代数

• $\{\mathscr{F}_\alpha\}_{\alpha\in I}$, $\mathscr{F}_\alpha$ 为 σ-代数 $\Rightarrow \bigcap_{\alpha\in I}\mathscr{F}_\alpha$ σ-代数 但 $\bigcup_{\alpha\in I}\mathscr{F}_\alpha$ 不一定.

记 $\bigvee_{\alpha\in I}\mathscr{F}_\alpha\triangleq\sigma(\bigcup_{\alpha\in I}\mathscr{F}_\alpha)$

§1.2 Dykin π-λ 定理.

Def. (π-类). 设 $\mathcal{A}\subset 2^\Omega$. 若 $\forall A_1,A_2\in\mathcal{A}$ 有 $A_1\cap A_2\in\mathcal{A}$, 则称 $\mathcal{A}$ 为一个 π-类.

RMK. • π类: 有限交封闭.

• π类典例: 代数, σ-代数; $\mathcal{A}=\{\mathbb{R}中开集\}$, $\mathcal{A}=\{(-\infty,b]: b\in\mathbb{R}\}$.

Def. (λ-类 / Dykin system). 设 $\mathcal{L}\subset 2^\Omega$, 若.

(i) $\Omega\in\mathcal{L}$

(ii) $A,B\in\mathcal{L}$ 且 $A\subset B$, 则 $B\setminus A\in\mathcal{L}$

(iii) $\forall A_n\uparrow\in\mathcal{L}\Rightarrow\bigcup_{i=1}^{\infty}A_i\in\mathcal{L}$.

则称 $\mathcal{L}$ 为一个 λ-类.

RMK. • (i) + (ii) $\Rightarrow$ λ-类补封闭

• σ代数 $\underset{\not\Leftarrow}{\Rightarrow}$ λ类

Thm. (等价 λ-类定义.) 设 $\mathcal{L}\subset 2^\Omega$. 若

(i)' $\Omega\in\mathcal{L}$

(ii)' $A\in\mathcal{L}\Rightarrow A^c\in\mathcal{L}$

(iii)' $A_n\in\mathcal{L}$ 且 $A_i\cap A_j=\phi\Rightarrow\bigcup_{i=1}^{\infty}A_i\in\mathcal{L}$ ($\Rightarrow\bigcup_{i=1}^{n}A_i\in\mathcal{L}$)

Thm ($\sigma=\pi+\lambda$) ($\sigma = a + m$ = 代数 + 单调类.) $\mathscr{F}$ 为 σ-代数 $\Leftrightarrow$ $\mathscr{F}$ 为 π类和 λ类 (($\Leftarrow$): 与 (ii)+(iii) $\Rightarrow$ (iii)' 类似)

Thm (Dykin π-λ类定理). 设 $\mathcal{A}$ 为 π-类, $\mathcal{L}$ 为 λ-类, 若 $\mathcal{A}\subset\mathcal{L}\Rightarrow\sigma(\mathcal{A})\subset\mathcal{L}$.

Cor. $\underset{\pi类}{\mathcal{A}}$ 为 π-类, 则 $\sigma(\mathcal{A})=\lambda(\mathcal{A})$ Pf: ($\Rightarrow$) $\underset{(\pi)}{\mathcal{A}}\subset\underset{(\lambda)}{\lambda(\mathcal{A})}\overset{}{\Rightarrow}\sigma(\mathcal{A})\subset\lambda(\mathcal{A})$ (π-λ定理). ($\Leftarrow$) $\sigma(\mathcal{A})$ 为 λ-类 $\therefore\lambda(\mathcal{A})\subset\sigma(\mathcal{A})$ #

$\boxed{\lambda(\mathcal{A})为\pi类} \Leftarrow \boxed{\lambda(\mathcal{A})=D_2} \Leftarrow \boxed{D_2为\lambda类(HW),\ \mathcal{A}\subset D_2} \Leftarrow \boxed{\forall B\in\mathcal{A},\forall C\in\lambda(\mathcal{A}),有B\cap C\in\lambda(\mathcal{A})}$

$\boxed{\mathcal{A}为\pi类} \Rightarrow \boxed{\forall B,C\in\mathcal{A},有B\cap C\in\lambda(\mathcal{A})} \Rightarrow \boxed{D_1为\lambda类(HW),\ \mathcal{A}\subset D_1} \Rightarrow \boxed{\lambda(\mathcal{A})=D_1} \Rightarrow \boxed{\forall B\in\mathcal{A},\forall C\in\lambda(\mathcal{A}),有B\cap C\in\lambda(\mathcal{A})}$

Ex. (函数形式的π-λ定理) 设 $\mathcal{A}\subset 2^{\Omega}$ 为π类且 $\Omega\in\mathcal{A}$.

设 $\mathcal{H}$ 是包含定义在Ω上某些实值函数的全体且

(i) $\forall A\in\mathcal{A}$, 有$\mathbb{1}_A\in\mathcal{H}$

(ii) $\forall f_1\in\mathcal{H}$, $f_2\in\mathcal{H}$, $c_1,c_2\in\mathbb{R}$ 有$c_1f_1+c_2f_2\in\mathcal{H}$. ($\mathcal{H}$为向量空间)

(iii) $\forall\, 0\le f_n\uparrow\in\mathcal{H}$ 且 $f_n\uparrow f$, 则 $f\in\mathcal{H}$.

则 $\mathcal{H}$ 包含所有 $\sigma(\mathcal{A})$-可测的函数.

(<u>RMK</u>: 称 $f:(\Omega,\sigma(\mathcal{A}))\to(\mathbb{R},\mathcal{B}_{\mathbb{R}})$ 为$\sigma(\mathcal{A})$-可测, 若 $\forall B\in\mathcal{B}_{\mathbb{R}}$, $f^{-1}(B)\in\sigma(\mathcal{A})$

其中 $f^{-1}(B)\triangleq\{w\in\Omega: f(w)\in B\}$.).

<u>pf</u>: 只需证 $\forall A\in\sigma(\mathcal{A})$, $\mathbb{1}_A\in\mathcal{H}$.

(若已证, 则由(ii),(iii) ⇒ $\mathcal{H}$ 包含所有$\sigma(\mathcal{A})$-可测的简单函数

⇒ 由$\forall$非负$\sigma(\mathcal{A})$可测f可被简单函数列$f_n\uparrow f$ 知 $f\in\mathcal{H}$.

⇒ $\forall\,\sigma(\mathcal{A})$可测$f$, $f=f^++f^-$, 知$f\in\mathcal{H}$.).

定义 $\mathcal{L}\triangleq\{A\subset\Omega: \mathbb{1}_A\in\mathcal{H}\}$, 由(i)知 $\mathcal{A}\subset\mathcal{L}$.

下证 $\mathcal{L}$ 为λ类: ① $\Omega\in\mathcal{L}$

② $A,B\in\mathcal{L}$, $A\subset B$.

$\because \mathbb{1}_{B\setminus A}=\mathbb{1}_B-\mathbb{1}_A\in\mathcal{H} \Rightarrow B\setminus A\in\mathcal{L}$

③ $\forall A_n\uparrow\in\mathcal{L}$.

$\because \mathbb{1}_{\bigcup_{n=1}^{\infty}A_n}\nwarrow\mathbb{1}_{A_n}$ $\therefore \mathbb{1}_{\cup A_n}\in\mathcal{H} \Rightarrow \cup A_n\in\mathcal{L}$

$\therefore$ 由π-λ定理 $\sigma(\mathcal{A})\subset\mathcal{L}$ $\therefore \forall A\in\sigma(\mathcal{A})$, $\mathbb{1}_A\in\mathcal{H}$. #.

<u>RMK</u> $\Omega=\mathbb{R}$, $\mathcal{A}=\{(-\infty,x]: x\in\mathbb{R}\}$, $\mathcal{H}$ 包含所有$\sigma(\mathcal{A})$-可测函数全体,

称为$\mathcal{B}_{\mathbb{R}}$-可测 或 Borel 函数.

Ex. (Dykin-乘法类定理) 设$\mathcal{H}$为包含Ω上某些实值有界函数的全体且

(i) $1\in\mathcal{H}$

(ii) $\mathcal{H}$为向量空间 ($\forall f_1,f_2\in\mathcal{H}$, $\forall c_1,c_2\in\mathbb{R}$, $c_1f_1+c_2f_2\in\mathcal{H}$.)

(iii) $0\le f_n\uparrow\in\mathcal{H}$, $f_n\uparrow f$, 则 $f\in\mathcal{H}$.

设 M 为一个包含所有常数的乘法类 ($\forall f_1,f_2\in M$, $f_1f_2\in M$) 且 $M\subset\mathcal{H}$.

则 $\mathcal{H}$ 包含所有 $\sigma(M)$-可测函数全体.

(<u>RMK</u>. $\sigma(M)\triangleq\sigma(\{f^{-1}(B): B\in\mathcal{B}_{\mathbb{R}}, f\in M\})$.)

<u>Pf</u>: 只需证 $\forall A\in\sigma(M)$, $\mathbb{1}_A\in\mathcal{H}$

定义 $\mathcal{L}\triangleq\{A\subset\Omega: \mathbb{1}_A\in\mathcal{H}\}$. 与左边一样, $\mathcal{L}$为λ-类.

构造π-类 $\mathcal{A}\subset\mathcal{L}$, s.t. $\sigma(\mathcal{A})=\sigma(M)$, 则由$\pi$-$\lambda$定理 $\sigma(M)\subset\mathcal{L}$.

$\mathcal{A}\triangleq\{\bigcap_{i=1}^{k}\{f_i>a_i\}: f_i\in M, a_i\in\mathbb{R}, k\in\mathbb{N}\}$ 【check: $\sigma(\mathcal{A})=\sigma(M)$, $\mathcal{A}$为π-类】

- 构造一列非负连续 $\varphi_n(x)\uparrow\mathbb{1}_{\{x>0\}}$, $x\in\mathbb{R}$.
- $\forall f_1,\cdots,f_k\in M$, $a_1,\cdots,a_k\in\mathbb{R}$, $k\in\mathbb{N}$,

定义: $F_n(w)\triangleq\prod_{i=1}^{k}\varphi_n(f_i(w)-a_i)$, $w\in\Omega$

则 F_n 非负$\uparrow\mathbb{1}_{\bigcap_{i=1}^{k}\{f_i>a_i\}}$, 只要 $F_n\in\mathcal{H}$, 则由(iii)知 $\mathcal{A}\subset\mathcal{L}$.

$\Rightarrow\sigma(\mathcal{A})=\sigma(M)\subset\mathcal{L}$.

- 定义 $M\triangleq\max_{1\le i\le k}\sup_{w\in\Omega}|f_i(w)-a_i|<+\infty$ (由$\mathcal{H}$有界)

$\varphi_n(x)\in C[-M,M]$, 由Weierestrass逼近定理, $\exists$多项式$P_l(x)$, s.t.

$\lim_{l\to\infty}\sup_{[-M,M]}|P_l(x)-\varphi_n(x)|=0$ (即 $P_l\overset{}{\underset{l\to+\infty}{\rightrightarrows}}\varphi_n$).

$\therefore \lim_{l\to\infty}\sup_{\Omega}\left|\prod_{i=1}^{k}P_l\circ(f_i-a_i)(w)-F_n(w)\right|=0$

且 $\prod_{i=1}^{k}P_l\circ(f_i-a_i)\in\mathcal{H}$ $\Downarrow_{l\to+\infty}$ F_n $\Big\}\Rightarrow F_n\in\mathcal{H}$.

【CLAIM: $f_n\in\mathcal{H}$, $f_n\rightrightarrows f$. 则 $f\in\mathcal{H}$.】 □

$f_{n_1},\ f_{n_2}+\frac{1}{2},\ f_{n_3}+\frac{1}{2}+\frac{1}{4}$

$\uparrow\rightrightarrows f+1\in\mathcal{H}$

$f\in\mathcal{H}$.

Pf $\sigma = a + m$:

(⇒) Obvious

(⇐) 只需证 $\mathcal{F}$ 可数并封闭，$\forall \{A_n\} \subset \mathcal{F}$.

$B_n \triangleq \bigcup_{i=1}^{n} A_n \in \mathcal{F}$ (由代数)

$\bigcup_{i=1}^{\infty} A_n = \bigcup_{i=1}^{\infty} B_n \in \mathcal{F}$ (由单调类). #.

Pf of Halmos: $m(\mathcal{A}) \triangleq \mathcal{A}$ 生成的(最小)单调类，则 $m(\mathcal{A}) \subset \mathcal{M}$

下证 $m(\mathcal{A})$ 为代数.

1° 有限交封闭 ⇐ (如果补封闭)，有限并封闭.

即证 $m(\mathcal{A})$ 为 π-类，证明与 π-λ thm 类似.

2° 补封闭

$D \triangleq \{B \in m(\mathcal{A}) : B^c \in m(\mathcal{A})\}$，check D 为单调类即可. $\mathcal{A} \subset D$ 且 (HW.)

∵ $\forall B \in \mathcal{A}$，$B^c \in \mathcal{A} \subset m(\mathcal{A})$ ∴ $\mathcal{A} \subset D$. #

RMK.2. $\forall \mathcal{F}$ 上有限测度 μ，$\forall A \in \mathcal{F}$.

$\mathbb{P}(A) \triangleq \frac{\mu(A)}{\mu(\Omega)}$ 为 $\mathcal{F}$ 上的一个概率测度.

3. $\forall \mathcal{F}$ 上 σ-有限测度 μ. $\{A_i\}$ 划分 Ω 且 $\mu(A_i) < +\infty$

$\mathbb{P}_i(A) \triangleq \frac{\mu(A \cap A_i)}{\mu(A_i)}$，$\forall A \in \mathcal{F}$. 为 $\mathcal{F}$ 上概率测度.

e.g. 1. $\forall x_1, \cdots, x_n \in \mathbb{R}$

$\mathbb{P} \triangleq \frac{1}{n} \sum_{i=1}^{n} \delta_{x_i}$ 为 $\mathcal{B}_{\mathbb{R}}$ 上的概率测度 (经验概率测度.)

2. $X_1, \cdots, X_n$ 为 n 个 r.v.s

$\mu \triangleq \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i}$

$\mu(w, A) \triangleq \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i(w)}(A)$. $A \in \mathcal{B}_{\mathbb{R}}$ 为随机概率测度.

$X_i(t, w)$ 为 n 个随机过程，$t \geq 0$，$w \in \Omega$

$\mu_t(w, A) \triangleq \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i(t,w)}(A)$ $\{\mu_t\}_{t \geq 0}$ 为概率测度值随机过程.

§1.3 单调类定理

Def.(单调类). 设 $\mathcal{M} \subset 2^{\Omega}$. 若

(i) $\forall A_n \uparrow \in \mathcal{M}$，$\bigcup_{i=1}^{\infty} A_n \in \mathcal{M}$

(ii) $\forall A_n \downarrow \in \mathcal{M}$，$\bigcap_{i=1}^{\infty} A_i \in \mathcal{M}$

则称 $\mathcal{M}$ 为单调类.

RMK. σ-代数 ⇒ 单调类.

λ-类 ⇒ 单调类.

Thm. ($\sigma = a + m$) 设 $\mathcal{F} \subset 2^{\Omega}$，则

$\mathcal{F}$ 为 σ-代数 ⇔ $\mathcal{F}$ 为代数 + 单调类.

Thm. (单调类定理, Halmos). 设 $\mathcal{A}$ 为代数，$\mathcal{M}$ 为单调类，则

$\mathcal{A} \subset \mathcal{M} \Rightarrow \sigma(\mathcal{A}) \subset \mathcal{M}$.

Cor. 设 $\mathcal{A}$ 为代数，则 $\sigma(\mathcal{A}) = m(\mathcal{A})$ 1° Halmos: $\sigma(\mathcal{A}) \subset m(\mathcal{A})$ 2° 由 $\sigma(\mathcal{A})$ 单调类 $m(\mathcal{A}) \subset \sigma(\mathcal{A})$.

§1.4 概率测度.

Def (概率测度) 可测空间 $(\Omega, \mathcal{F})$，$\mathbb{P} : \mathcal{F} \to [0,1]$ (集函数.) 若

(i) $\mathbb{P}(\phi) = 0$，$\mathbb{P}(\Omega) = 1$.

(ii) $\forall \{A_n\} \subset \mathcal{F}$，$A_i \cap A_j = \phi \, (i \neq j)$，

$\mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$ (可列可加性/σ-可加性.)

则称 $\mathbb{P}$ 为定义在事件域 $\mathcal{F}$ 上的一个概率测度.

RMK1. 若 $\mu : \mathcal{F} \to [0, +\infty]$ 满足 (i)(ii)，则称 μ 为 $\mathcal{F}$ 的一个测度

进一步，若 $\mu(\Omega) < +\infty$，称 μ 为 $\mathcal{F}$ 上的有限测度.

若 $\mu(\Omega) = +\infty$，但 $\exists \, \Omega$ 上的一个划分 $\{A_i\}$，s.t. $\Omega = \bigcup_{i=1}^{\infty} A_i$，$A_i \subset \mathcal{F}$，$A_i \cap A_j = \phi_{(i \neq j)}$

且 $\mu(A_i) < +\infty$，则称 μ 为 $\mathcal{F}$ 上的一个 σ-有限测度.

(e.g. $\Omega = \mathbb{R}$，$\mathcal{F} = \mathcal{B}_{\mathbb{R}}$，$\mu$: $\mathcal{B}_{\mathbb{R}}$ 上的 Lebesgue 测度.)

若取 $\Omega = [0,1]$，μ 为 $\mathcal{B}_{[0,1]}$ 上的有限测度 且为概度测度).

e.g. $\Omega=\{w_1,\cdots,w_n,\cdots\}$，$\mathcal{F}=2^\Omega$. 设 $p(w):\Omega\to\mathbb{R}_+$ 满足 $\sum_\Omega p(w)=1$.

定义 $\forall A\in\mathcal{F}$，$\mathbb{P}(A)=\sum_{w\in A}p(w)$ 为 $\mathcal{F}$ 上一个概率测度.

特别①若 $\Omega=\{w_1,\cdots,w_n\}$，$\mathcal{F}=2^\Omega$，$\mathbb{P}(A)=\frac{|A|}{n}$ 为概率测度.（等可能性）

称 $(\Omega,2^\Omega,\mathbb{P})$ 为古典概型的概率空间.

②几何概型. $\Omega=\{(x,y):x^2+y^2\le R^2\}$

$\mathcal{F}=\{\Omega$ 中具有面积的区域 $O\}$.

$\mathbb{P}(O)=\frac{|O|}{\pi R^2}$，$O\in\mathcal{F}$.（等可能性）

为 $\mathcal{F}$ 上的一个概率测度.

e.g. 无穷次掷硬币的随机实验.

$\Omega=\{(w_1,\cdots,w_n,\cdots):w_i\in\{0,1\},i\ge1\}$

$\mathcal{F}=2^\Omega$

$\Omega_n\triangleq\{(w_1,\cdots,w_n):w_i\in\{0,1\},i\ge1\}$，则 $|\Omega_n|=2^n$.

定义 $\mathbb{P}_n(A)=\frac{|A|}{2^n}$，$\forall A\in2^{\Omega_n}$

定义 $\mathcal{F}_n\triangleq\{A\subset\Omega:\exists B\in2^{\Omega_n},s.t.\ A=\{w=(w_i)_{i\ge1},(w_1,\cdots,w_n)\in B\}.\}$

• Q: (i) $\mathcal{F}_n$ 是 σ-代数吗？

(ii) $\{\mathcal{F}_n\}_{n\ge1}$ 是一个过滤(单增)吗？

$\forall A\in\mathcal{F}_n$，$\exists B\in2^{\Omega_n}$, s.t. $A=\{w:(w_1,\cdots,w_n)\in B\}$，定义

$\mathbb{P}(A)\triangleq\mathbb{P}_n(B)$

• Q: $\mathbb{P}$ 为 $\mathcal{F}_n$ 上概率测度吗？

于是建立了 $\bigcup_{n=1}^\infty\mathcal{F}_n$ 上的一个集函数 $\mathbb{P}$

• CLAIM: $\bigcup\mathcal{F}_n$ 是代数但不是 σ-代数.

• Q: 怎么将 $\mathbb{P}$ 从 $\mathcal{F}_\infty\triangleq\bigcup\mathcal{F}_n$ 上扩展到 $\sigma(\mathcal{F}_\infty)$？(去年期末 check $\mathbb{P}$ 在 $\mathcal{F}_\infty$ 上可列可加)

Thm (Carathéodory 延拓) 设 $\mathcal{A}$ 为代数，$\mu_0:\mathcal{A}\to[0,+\infty]$ 是可列可加的集函数

则 $\exists(\Omega,\sigma(\mathcal{A}))$ 上的测度 μ, s.t. $\mu|_{\mathcal{A}}=\mu_0$. 若 $\mu_0(\Omega)<+\infty$，则 μ 唯一.

RMK: μ_0 在 $\mathcal{A}$(代数)上可列可加性的定义: $\forall A_n\in\mathcal{A}$, $A_i\cap A_j=\phi$, 且 $\bigcup_{n=1}^\infty A_n\in\mathcal{A}$，有 $\mu_0(\bigcup A_n)=\sum\mu_0(A_n)$

（概率测度性质.）

Prop (1) $\forall A,B\in\mathcal{F}$，$P(A\cup B)=P(A)+P(B)-P(A\cap B)$

(容斥公式)

(2) $\forall A_1,\cdots,A_n\in\mathcal{F}$，$P(\bigcup_{i=1}^n A_i)=\sum_{i=1}^n(-1)^{i-1}\sum_{\substack{J\subset\{1,\cdots,n\}\\|J|=i}}P(\bigcap_{j\in J}A_j)$（归纳证.）

(3) $P(\bigcap_{i=1}^n A_i)=\sum_{i=1}^n(-1)^{i-1}\sum_{|J|=i}P(\bigcup_{j\in J}A_j)$

∵ $P(\bigcup A_i^c)=\sum_{i=1}^n(-1)^{i-1}\sum_{|J|=i}P(\bigcap_{j\in J}A_j^c)$

$=\sum_{i=1}^n(-1)^{i-1}\sum_{|J|=i}[1-P(\bigcup_{j\in J}A_j)]=1-\sum_{i=1}^n(-1)^{i-1}\sum_{|J|=i}P(\bigcup_{j\in J}A_j)$

(4) (Bonferroni) $\forall A_1,\cdots,A_n\in\mathcal{F}$，$P(\bigcup_{i=1}^n A_i)\begin{cases}\le\sum_{i=1}^m(-1)^{i-1}\sum_{|J|=i}P(\bigcap_{j\in J}A_j), & m\text{奇}\\ \ge\sum_{i=1}^m(-1)^{i-1}\sum_{|J|=i}P(\bigcap_{j\in J}A_j), & m\text{偶}.\end{cases}$

(5) (Boole) $\forall A_i\in\mathcal{F}$，$P(\bigcup_{i=1}^n A_i)\le\sum_{i=1}^n P(A_i)$

进一步，有 $\overline{\lim_{n\to\infty}}P(\bigcup_{i=1}^n A_i)\le\sum_{i=1}^\infty P(A_i)$

(6) (上.下连续性). 下: $A_n\in\mathcal{F}\uparrow\Rightarrow P(\bigcup_{i=1}^\infty A_i)=\lim_{n\to\infty}P(A_n)$

上: $A_n\in\mathcal{F}\downarrow\Rightarrow P(\bigcap_{i=1}^\infty A_i)=\lim_{n\to\infty}P(A_n)$

∵ 下: $B_n\triangleq A_{n+1}\setminus A_n$ 互斥，$P(\bigcup A_i)=P(A_n\cup\bigcup_{i=n}^\infty B_i)=P(A_n)+\sum_{i=n}^\infty P(B_i)$

再由 $P(\bigcup A_i)=P(A_1)+\sum_{i=1}^\infty P(B_i)\Rightarrow\sum P(B_i)<\infty\Rightarrow\sum_{i=n}^\infty P(B_i)\to0$ 得证.

(7) 设 $\mathcal{F}$ 为 σ-代数，$\mu:\mathcal{F}\to[0,\infty]$ 为有限可加集函数，$\mu(\phi)=0$.

若 $\forall H_n\in\mathcal{F}\downarrow\phi$，有 $\mu(H_n)\downarrow0$，则 μ 在 $\mathcal{F}$ 上可列可加.

∵ $H_n:=\bigcup_{i=n+1}^\infty A_i$，$A_i\cap A_j=\phi$，则 $\mu(\bigcup A_i)=\mu(\bigcup_{i=1}^n A_i\cup H_n)=\sum_{i=1}^n\mu(A_i)+\mu(H_n)\to\sum_{i=1}^\infty\mu(A_i)$.

RMK. 若 $\mathcal{F}$ 为代数，此结果也成立.

§1.5 Borel-Cantelli 引理.

Def. $\overline{\lim_{n\to\infty}}A_n=\bigcap_{m=1}^\infty\bigcup_{n=m}^\infty A_n$，$\underline{\lim_{n\to\infty}}A_n=\bigcup_{m=1}^\infty\bigcap_{n=m}^\infty A_n$

RMK. $\overline{\lim}A_n,\underline{\lim}A_n\in\mathcal{F}$. $\underline{\lim}A_n\subset\overline{\lim}A_n$

$\overline{\lim}A_n=\{w\in\Omega:\forall m\ge1,\exists n(w)\ge m,s.t.\ w\in A_{n(w)}\}$

$=\{w\in\Omega:w\in A_n$ for infinitely many $n\}$

$=\{A_n,\text{ i.o.}\}$.

$\underline{\lim}A_n=\{w\in\Omega:\exists m(w)\ge1,\forall n\ge m(w),w\in A_n\}$

$=\{w\in\Omega:w\in A_n$ for large $n\}$

$=\{A_n,\text{ e.v.}\}$

e.g. $\Omega=[0,1]$, $\mathcal{A}=\{\bigcup_{i=1}^{n}(a_i,b_i],\ 0\le a_1<b_1<\cdots<a_n<b_n\le 1\}$ 为一代数 不为σ-代数

定义 $\mu_0(\bigcup_{i=1}^{n}(a_i,b_i])\triangleq\sum_{i=1}^{n}(b_i-a_i)$，则 μ_0 为 $\mathcal{A}$ 上有限可加集函数，$\mu_0(\phi)=0$

问：能否将 μ_0 从 $\mathcal{A}$ 延拓到 $\sigma(\mathcal{A})=\mathcal{B}_{[0,1]}$ 上的一个测度？

Sol：若用 Carathéodory 延拓 thm，则要 μ_0 在 $\mathcal{A}$ 上可列可加.

只须证：$\forall\{A_n\}$，$\bigcup_{n=1}^{\infty}A_n\in\mathcal{A}\Rightarrow\mu_0(H_n)\downarrow 0$，其中 $H_n\triangleq\bigcup_{i=1}^{\infty}A_i\setminus\bigcup_{i=1}^{n}A_i\downarrow\phi$

反证：若 $\exists\,\varepsilon>0$, s.t. $\mu_0(H_n)\ge 2\varepsilon$.

由 $\mathcal{A}$ 和 μ_0 的定义，$\forall l\ge 1$，$\exists J_l\in\mathcal{A}$，$\overline{J_l}\subset H_l$，s.t. $\mu_0(H_l\setminus J_l)\le\varepsilon 2^{-l}$

$\therefore\ \mu_0(\bigcap_{l\le n}J_l)=\mu_0(H_n)-\mu_0(H_n\setminus\bigcap_{l\le n}J_l)$

$\because\ \forall l\le n,\ H_n\subset H_l$

$\therefore\ H_n\setminus\bigcap_{l\le n}J_l=\bigcup_{l\le n}(H_n\setminus J_l)\subset\bigcup_{l\le n}(H_l\setminus J_l)$

$\therefore\ \mu_0(H_n\setminus\bigcap_{l\le n}J_l)\le\sum_{l\le n}\mu_0(H_l\setminus J_l)\le\varepsilon.$

$\therefore\ \mu_0(\bigcap_{l\le n}J_l)\ge\varepsilon>0.$

定义 $C_n\triangleq\bigcap_{l\le n}\overline{J_l}\neq\phi$，则 $C_n\downarrow\bigcap_{l\ge 1}\overline{J_l}$，

由 C_n 紧知 $\bigcap_{l\ge 1}\overline{J_l}\neq\phi$，与 $\bigcap_{l\ge 1}\overline{J_l}\subset\bigcap_{l\ge 1}H_l=\phi$ 矛盾！ #.

Pf of Fatou: $P(\underline{\lim}A_n)=P(\bigcup_{m=1}^{\infty}\underbrace{\bigcap_{n=m}^{\infty}A_n}_{G_m})=\lim_{m\to\infty}P(G_m)\le\lim_{m\to\infty}\inf_{n\ge m}P(A_n)=\underline{\lim}P(A_n)$

$P(\overline{\lim}A_n)=1-P(\underline{\lim}A_n^c)\ge 1-\underline{\lim}(1-P(A_n))=\overline{\lim}P(A_n)$ #

Pf of B-C I: $P(\overline{\lim}A_n)=\lim_{m\to\infty}P(\bigcup_{n=m}^{\infty}A_n)\overset{\text{Boole}}{\le}\lim_{m\to\infty}\sum_{n=m}^{\infty}P(A_n)=0.$ #.

Pf of B-C II: 只须 $P(\underline{\lim}A_n^c)=0$

$=P(\bigcup_{m=1}^{\infty}\bigcap_{n=m}^{\infty}A_n^c)=\lim_{m\to\infty}P(\bigcap_{n=m}^{\infty}A_n^c)=\lim_{m\to\infty}\prod_{n=m}^{\infty}(1-P(A_n))\le\lim_{m\to\infty}e^{-\sum_{n=m}^{\infty}P(A_n)}=0.$ #

· $1_{\underline{\lim}A_n}(\omega)=\underline{\lim}\,1_{A_n}(\omega)$，$1_{\overline{\lim}A_n}(\omega)=\overline{\lim}\,1_{A_n}(\omega)$

Lem (Fatou) $P(\underline{\lim}A_n)\le\underline{\lim}P(A_n)\le\overline{\lim}P(A_n)\le P(\overline{\lim}A_n)$

Lem (Borel-Cantelli I) 设 $\{A_n\}\subset\mathcal{F}$，$\sum_{n=1}^{\infty}P(A_n)<+\infty$，则 $P(\overline{\lim}A_n)=0$

Rmk $X_n\xrightarrow{a.s.}X\Leftrightarrow P(\lim X_n=X)=1$

【 $\forall\omega\in\{\omega\in\Omega:\lim_{n\to\infty}X_n(\omega)=X(\omega)\}$，$\lim_{n\to\infty}X_n(\omega)=X(\omega)$，即

$\forall\varepsilon>0$，$\exists m(\omega)\ge 1$，s.t. $\forall n\ge m(\omega)$，$|X_n(\omega)-X(\omega)|<\varepsilon$.

$\Leftrightarrow\omega\in\bigcup_{m\ge 1}\bigcap_{n\ge m}\{\omega\in\Omega:|X_n(\omega)-X(\omega)|\le\varepsilon\}$. 定义 $A_n\triangleq\{\omega\in\Omega:|X_n(\omega)-X(\omega)|>\varepsilon\}$

$\therefore P\{\omega\in\Omega:\lim X_n(\omega)=X(\omega)\}=P(\bigcup_{m\ge 1}\bigcap_{n\ge m}A_n^c)=1\Leftrightarrow P(\overline{\lim}A_n)=0,\ \forall\varepsilon>0.$

$\Leftrightarrow P(|X_n-X|>\varepsilon,\ i.o.)=0,\ \forall\varepsilon>0.$

$\sum_{n=1}^{\infty}P(\{|X_n-X|>\varepsilon\})<+\infty,\ \forall\varepsilon>0\overset{B\text{-}C\text{-}I}{\Rightarrow}X_n\xrightarrow{a.s.}X.$ 】

Lem (Borel-Cantelli II) 设 $\{A_n\}\in\mathcal{F}$，且相互独立. 若 $\sum_{n=1}^{\infty}P(A_n)=+\infty$，

则 $P(\overline{\lim}A_n)=1$

Thm. (Kolmogorov 0-1 律.) 设 $A_n\in\mathcal{F}$ 且相互独立，则 $\mathbb{P}(\overline{\lim}A_n)=0$ 或 1

§1.6. 概率空间完备化.

Def. (零事件. null event) $(\Omega,\mathcal{F},\mathbb{P})$，称 N 为一个零事件，若 $N\in\mathcal{F}$，$\mathbb{P}(N)=0$.

记 $\mathcal{N}\triangleq\{N\in\mathcal{F}:\mathbb{P}(N)=0\}$ 为所有零事件全体

Rmk. $\forall\{N_j\}_{j\ge 1}\subset\mathcal{N}$，$\bigcup_{j=1}^{\infty}N_j\in\mathcal{N}$，$\bigcap_{j=1}^{\infty}N_j\in\mathcal{N}$

Def. (完备概率空间) $(\Omega,\mathcal{F},\mathbb{P})$：$\forall N\in\mathcal{N}$ 且 $B\subset N$，有 $B\in\mathcal{N}$，则称为一个完备的概率空间.

Thm. (概率空间完备化) $\forall(\Omega,\mathcal{F},\mathbb{P})$，$\exists(\Omega,\overline{\mathcal{F}},\overline{\mathbb{P}})$，s.t.

$\mathcal{F}\subset\overline{\mathcal{F}}$，$\overline{\mathbb{P}}|_{\mathcal{F}}=\mathbb{P}$

Pf: $\overline{\mathcal{F}}\triangleq\mathcal{F}\cup\overline{\mathcal{N}}=\{E=A\cup B;\ A\in\mathcal{F},\ B\subset N\in\mathcal{N}\}$，其中 $\overline{\mathcal{N}}=\{B\subset\Omega:B\subset N\in\mathcal{N}\}$

Check $\overline{\mathcal{F}}$ 为 σ-alg: ① $\Omega\in\overline{\mathcal{F}}$：$\Omega=\Omega\cup\phi$

② $\forall E\in\overline{\mathcal{F}}$，$E^c=A^c\cap B^c=\underbrace{(A^c\cap N^c)}_{\in\mathcal{F}}\cup\underbrace{(A^c\cap B^c\cap N)}_{\in\overline{\mathcal{N}}}\in\overline{\mathcal{F}}$

③ $\forall E_j=A_j\cup B_j\in\overline{\mathcal{F}}$，

$\cup E_j=(\cup A_j)\cup(\cup B_j)\in\mathcal{F}\cup\overline{\mathcal{N}}=\overline{\mathcal{F}}$

Ex. $\mathcal{H} \triangleq \{F \subset \Omega, \exists G \in \mathcal{F}, \text{s.t. } G \Delta F \in \overline{\mathcal{N}}\}$

证明：(i) $\mathcal{H}$ 为 σ-代数

(ii) $\mathcal{H} = \overline{\mathcal{F}} = \mathcal{F} \cup \overline{\mathcal{N}}$

Sol: (i) ① $\Omega \in \mathcal{H}$: $\Omega \Delta \Omega = \phi \in \overline{\mathcal{N}}$

② $\left.\begin{array}{l} F \in \mathcal{H},\ F^c \Delta G^c = F \Delta G \in \overline{\mathcal{N}} \\ G^c \in \mathcal{F} \end{array}\right\} \Rightarrow F^c \in \mathcal{H}$

③ $\left.\begin{array}{l} \{F_j\} \subset \mathcal{H},\ (\bigcup F_j) \Delta (\bigcup G_j) \subset \bigcup (F_j \Delta G_j) \in \overline{\mathcal{N}} \\ \bigcup G_j \in \mathcal{F} \end{array}\right\} \Rightarrow \bigcup F_j \in \mathcal{H}$

(ii) $\mathcal{F} \subset \mathcal{H}$, $\overline{\mathcal{N}} \subset \mathcal{H}$ 显然.

$\therefore \overline{\mathcal{F}} \subset \mathcal{H}$.

下证 $\mathcal{H} \subset \overline{\mathcal{F}}$: $\forall F \in \mathcal{H}$, $F \Delta G = N \Leftrightarrow F = G \Delta N$

$\therefore \mathcal{H} = \{G \Delta N : G \in \mathcal{F}, N \in \overline{\mathcal{N}}\}$

$\because G \Delta N = \underbrace{(G \cap A^c)}_{\in \mathcal{F}} \cup \underbrace{(G \Delta N) \cap A}_{\mathcal{N}}$，其中 $A \in \mathcal{N}$

$\therefore \mathcal{H} \subset \overline{\mathcal{F}}$ #.

e.g. 1° $(\Omega, \mathcal{F}) = (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$，则 $X(w) = w : (\mathbb{R}, \mathcal{B}_{\mathbb{R}}) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ 为 r.v.

$(\Omega, \mathcal{F}) = (\mathbb{R}, \sigma(\{[a,b] : a<b \in \mathbb{R}\}))$，则 $X(w) = w$ 不是 r.v.

2. (1) 示性 r.v. $\forall A \in \mathcal{F}$, $\mathbb{1}_A(w) \triangleq \begin{cases} 1, & w \in A \\ 0, & w \notin A \end{cases} : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ 称为示性 r.v.

$\forall B \in \mathcal{B}_{\mathbb{R}}$ 且 $\mathbb{1}_A^{-1}(B) = \begin{cases} \Omega, & 0,1 \in B \\ A^c, & 0 \in B, 1 \notin B \\ A, & 1 \in B, 0 \notin B \\ \phi, & 0,1 \notin B \end{cases}$ $\Rightarrow \{\mathbb{1}_A^{-1}(B), B \in \mathcal{B}_{\mathbb{R}}\} = \{\phi, A, A^c, \Omega\} \subset \mathcal{F}$.

(2) 简单 r.v. $\forall \{A_i\}_1^n \subset \mathcal{F}$ 为 Ω 的一个划分. $X(w) \triangleq \sum_{i=1}^n a_i \mathbb{1}_{A_i}(w)$, $w \in \Omega$. 称为简单 r.v.

则 $\{X^{-1}(B) : B \in \mathcal{B}_{\mathbb{R}}\} = \sigma(\{A_1, \cdots, A_n\}) = \{\bigcup_{i \in I} A_i : I \subset \{1, \cdots, n\}\} \subset \mathcal{F}$，其中 $\bigcup_{i \in \phi} \triangleq \phi$.

$\overline{\mathbb{P}}(E) \triangleq \mathbb{P}(A)$ check: ① 良定: $E = A_1 \cup B_1 = A_2 \cup B_2$, 则 $\mathbb{P}(A_1) = \mathbb{P}(A_2)$

$\because A_1 \subset A_2 \cup B_2 \subset A_2 \cup N_2 \Rightarrow \mathbb{P}(A_1) \le \mathbb{P}(A_2) + \mathbb{P}(N_2) = \mathbb{P}(A_2)$

同理 $\mathbb{P}(A_2) \le \mathbb{P}(A_1)$

② $\overline{\mathbb{P}}$ 为 $\overline{\mathcal{F}}$ 上概率测度. #.

§1.7 随机变量

Def. (r.v.) 设 $X(w) : (\Omega, \mathcal{F}) \to (S, \mathcal{S})$ 为样本空间上的函数.

若 $\forall B \in \mathcal{S}$, $X^{-1}(B) \triangleq \{w \in \Omega : X(w) \in B\} \in \mathcal{F}$,

则称 $X(w)$ 为一个随机变量，记 $X \in \mathcal{F}$.

Rmk. $\mathcal{F} = 2^{\Omega}$ 时，任何 Ω 上函数都为 r.v.

Prop. r.v. $X(w) : (\Omega, \mathcal{F}) \to (S, \mathcal{S})$ 且 $f : (S, \mathcal{S}) \to (T, \mathcal{T})$ 可测函数 (i.e., $\forall B \in \mathcal{T}$, $f^{-1}(B) \in \mathcal{S}$.) 则 $f \circ X(w) : (\Omega, \mathcal{F}) \to (T, \mathcal{T})$ 也为一个 r.v.

$\because (f \circ X)^{-1}(B) = \{w \in \Omega : f(X(w)) \in B\} = \{w \in \Omega : X(w) \in f^{-1}(B) \in \mathcal{S}\} \overset{\text{r.v.}}{\in} \mathcal{F}$.

Lem. 设 $X(w) : (\Omega, \mathcal{F}) \to (S, \mathcal{S})$ 为样本空间上的函数，则

$\{X^{-1}(B) : B \in \mathcal{S}\}$ 为 σ-代数. 【Hint: $(X^{-1}(B))^c = X^{-1}(B^c)$, $X^{-1}(\bigcup_{\alpha \in I} B_\alpha) = \bigcup_{\alpha \in I} X^{-1}(B_\alpha)$

$X^{-1}(\bigcap_{\alpha \in I} B_\alpha) = \bigcap_{\alpha \in I} X^{-1}(B_\alpha)$.】

Rmk. 记 $\sigma(X) \triangleq \{X^{-1}(B) : B \in \mathcal{S}\}$.

Lem. 设 $X(w) : (\Omega, \mathcal{F}) \to (S, \mathcal{S})$，则 $X \in \mathcal{F} \Leftrightarrow \sigma(X) \subset \mathcal{F}$

Rmk. 1. 若 $X \in \mathcal{F}$，$\sigma(X)$ 为一个子事件域.

2. 当 $(S, \mathcal{S}) = (\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d})$，$X(w) = (X_1(w), \cdots, X_d(w))$，则 $\sigma(X) = \sigma((X_1 \cdots X_d)) = \sigma(\bigcup_{i=1}^d \sigma(X_i))$

当 $(S, \mathcal{S}) = (\mathbb{R}^\infty, \mathcal{B}_\infty)$，$X(w) = (X_1(w), \cdots\cdots)$,

则 $\sigma(X) = \sigma(\bigcup_{d=1}^\infty \sigma(X_1 \cdots X_d))$

3. $\sigma(X)$ 是使 $X(w)$ 为个 r.v. 的最小事件域.

Lem. 设 r.v. $X(w) : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$，则 $\exists$ r.v.s $\{X_n\}_{n \ge 1}$ 简单, s.t. $X_n(w) \to X(w)$, $\forall w \in \Omega, n \to \infty$.

Pf: $f_n(x) \triangleq n \mathbb{1}_{(n, +\infty)}(x) + \sum_{k=0}^{n2^n - 1} \frac{k}{2^n} \cdot \mathbb{1}_{(\frac{k}{2^n}, \frac{k+1}{2^n}]}(x)$, $x \ge 0$，则 $f_n(x) \nearrow x$, $\forall x \ge 0$.

$X(w) = X^+(w) - X^-(w)$，则 $\exists\ Y_n(w) \triangleq f_n(X^+(w)) \nearrow (X^+(w))$ $\Rightarrow X_n(w) \triangleq f_n(X^+(w)) - f_n(X^-(w)) \to X(w)$

$Z_n(w) \triangleq f_n(X^-(w)) \nearrow (X^-(w))$

#

E.g. 设 r.v.s $X_n(\omega):(\Omega,\mathcal{F})\to(\overline{\mathbb{R}},\mathcal{B}_{\overline{\mathbb{R}}})$，其中 $\overline{\mathbb{R}}=[-\infty,+\infty]$，

则 $\inf\limits_{n\ge1}X_n$，$\sup\limits_{n\ge1}X_n$，$\overline{\lim\limits_{n\to\infty}}X_n$，$\underline{\lim\limits_{n\to\infty}}X_n$ 也为 r.v.s.

Pf：取 $\mathcal{A}\triangleq\{[-\infty,x):x\in\mathbb{R}\}$，则 $\mathcal{B}_{\overline{\mathbb{R}}}=\sigma(\mathcal{A})$

$\{\omega\in\Omega:\inf\limits_{n\ge1}X_n(\omega)<x\}\xlongequal{}\bigcup\limits_{n=1}^{\infty}X_n^{-1}([-\infty,x))\in\mathcal{F}$.

$\therefore\inf\limits_{n\ge1}X_n\in\mathcal{F}$.

$\sup\limits_{n\ge1}X_n=-\inf\limits_{n\ge1}(-X_n)\in\mathcal{F}$

$\overline{\lim}X_n=\lim\limits_{m\to\infty}\sup\limits_{n\ge m}X_n=\inf\limits_{m\ge1}(\sup\limits_{n\ge m}X_n)\in\mathcal{F}$. #

HW. 设 (X,Y) 为 $(\Omega,\mathcal{F},\mathbb{P})$ 下二维 r.v.，则

$\exists$ 可测 f, s.t. $Y=f(X)\iff\sigma(Y)\subset\sigma(X)$.

注：设 Y 为 $\sigma(X)$-可测 $(Y\in\sigma(X))\Rightarrow\exists$ 可测 $f:\mathbb{R}\to\mathbb{R}$, s.t. $Y=f(X)$.

$Y:(\Omega,\sigma(X))\to(\mathbb{R},\mathcal{B}_{\mathbb{R}})$

称 f 为回归函数.

Lem. 设函数 $X(\omega):(\Omega,\mathcal{F})\to(S,\mathcal{S}=\sigma(\mathcal{A}))$，$\mathcal{A}\subset2^S$. 若 $\forall B\in\mathcal{A}$，$X^{-1}(B)\in\mathcal{F}$，

则 $X\in\mathcal{F}$.

Pf：只须证 $\forall B\in\sigma(\mathcal{A})$，$X^{-1}(B)\in\mathcal{F}$. 定义 $\mathcal{M}\triangleq\{B\subset S:X^{-1}(B)\in\mathcal{F}\}$，则 $\mathcal{A}\subset\mathcal{M}$

check $\mathcal{M}$ 为 σ-代数，故 $\sigma(\mathcal{A})\subset\mathcal{M}$. #

Rmk. 当 $(S,\mathcal{S})=(\mathbb{R},\mathcal{B}_{\mathbb{R}})$，$\mathcal{A}=\{(-\infty,x]:x\in\mathbb{R}\}$，则 $\mathcal{B}_{\mathbb{R}}=\sigma(\mathcal{A})$

$\therefore\forall x\in\mathbb{R}$，$X^{-1}((-\infty,x])\in\mathcal{F}\Rightarrow X\in\mathcal{F}$.

或 $X^{-1}((-\infty,x))\in\mathcal{F}\Rightarrow X\in\mathcal{F}$.

Lem. 设 r.v. $X(\omega):(\Omega,\mathcal{F})\to(S,\sigma(\mathcal{A}))$，$\mathcal{A}\subset2^S$，则 $\sigma(X)=\sigma(\{X^{-1}(B):B\in\mathcal{A}\})$

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

Pf：记 $\mathcal{G}\triangleq\sigma(\{X^{-1}(B):B\in\mathcal{A}\})$，

则 $\{X^{-1}(B):B\in\mathcal{A}\}\subset\{X^{-1}(B):B\in\sigma(\mathcal{A})\}=\sigma(X)\Rightarrow\mathcal{G}\subset\sigma(X)$.

$\because X\in\mathcal{F}\Rightarrow\sigma(X)\subset\mathcal{F}$ $\therefore\mathcal{G}\subset\sigma(X)\subset\mathcal{F}$.

下证 $X\in\mathcal{G}$. 即 $\forall B\in\sigma(\mathcal{A})$，$X^{-1}(B)\in\mathcal{G}$.

令 $\mathcal{M}\triangleq\{B\subset S:X^{-1}(B)\in\mathcal{G}\}$，则 $\mathcal{M}$ 为 σ-代数且 $\mathcal{A}\subset\mathcal{M}\Rightarrow\sigma(\mathcal{A})\subset\mathcal{M}$.

$\therefore$ $\mathcal{G}$ 为使 $X(\omega)$ 为 r.v. 的一个子事件域 $\Rightarrow\mathcal{G}=\sigma(X)$.

$\sigma(X)$ 为使 $X(\omega)$ 为 r.v. 的最小事件域，$\mathcal{G}\subset\sigma(X)$ #

Rmk. 当 $(S,\sigma(\mathcal{A}))=(\mathbb{R},\mathcal{B}_{\mathbb{R}})$，则 $X\in\mathcal{F}\Rightarrow\sigma(X)=\sigma(\{X^{-1}((-\infty,x]):x\in\mathbb{R}\})$

$(S,\sigma(\mathcal{A}))=(\mathbb{R}^d,\mathcal{B}_{\mathbb{R}^d})$，则 $X\in\mathcal{F}\Rightarrow\sigma(X)=\sigma(\{X^{-1}(\prod\limits_{i=1}^{d}(-\infty,x_i]):x_i\in\mathbb{R}\})$

$=\sigma(\{\bigcap\limits_{i=1}^{d}X_i^{-1}((-\infty,x_i]):x_i\in\mathbb{R}\})$.

Pf of Lem (同分布等价定义:分布函数定义): 只须证 $\forall B\in\mathcal{B}_{\mathbb{R}}$, $P_X(B)=P_Y(B)$.

已知 $P_X((-\infty,x])=P_Y((-\infty,x]),\ \forall x\in\mathbb{R}$.

令 $\mathcal{A}\triangleq\{(-\infty,x]:x\in\mathbb{R}\}$, 则 $\mathcal{A}$ 为 π 类, $\sigma(\mathcal{A})=\mathcal{B}_{\mathbb{R}}$

$\mathcal{L}\triangleq\{B\in\mathcal{B}_{\mathbb{R}}:P_X(B)=P_Y(B)\}$, 则 $\mathcal{A}\subset\mathcal{L}$

$\because \mathbb{R}\in\mathcal{L}\ (P_X(\mathbb{R})=P_Y(\mathbb{R})=1)$

$\forall A\subset B\in\mathcal{L}$, $P_X(B\setminus A)=P_X(B)-P_X(A)=P_Y(B)-P_Y(A)=P_Y(B\setminus A)$

$\Rightarrow B\setminus A\in\mathcal{L}$.

$\forall A_n\uparrow\in\mathcal{L}$, $A_n\uparrow\bigcup_n A_n\Rightarrow P_X(\bigcup_n A_n)=\lim_n P_X(A_n)=\lim_n P_Y(A_n)=P_Y(\bigcup_n A_n)$

$\Rightarrow\bigcup_n A_n\in\mathcal{L}$

$\therefore \mathcal{L}$ 为 λ 类

由 π-λ 定理

$\therefore \sigma(\mathcal{A})\subset\mathcal{L}\subset\mathcal{B}_{\mathbb{R}}\ \Rightarrow\mathcal{L}=\mathcal{B}_{\mathbb{R}}$ #.

第二章 分布与积分(数学期望)

Def. (分布) 设 r.v. $X(\omega):(\Omega,\mathcal{F})\to(S,\mathcal{S})$, 称 $P_X:\mathcal{S}\to[0,1]$ 为 r.v. X 的分布

$$B\mapsto P_X(B)\triangleq P(X^{-1}(B))$$

若 $(S,\mathcal{S})=(\mathbb{R},\mathcal{B}_{\mathbb{R}})$. 称 $F_X:\mathbb{R}\to[0,1]$ 为 r.v. X 的分布函数.

$$x\mapsto F_X(x)\triangleq P_X((-\infty,x])$$

Rmk 1. $(S,\mathcal{S},P_X)$ 为一个概率空间.

$(\Omega,\mathcal{F},\mathbb{P})\xRightarrow{def}$ r.v. $X(\omega):(\Omega,\mathcal{F})\to(S,\mathcal{S})\xRightarrow{def 分布}(S,\mathcal{S},P_X)$ 概率空间.

2 若 $(S,\mathcal{S})=(\mathbb{R}^d,\mathcal{B}_{\mathbb{R}^d})$, $F_X(x_1,\cdots,x_d)\triangleq P_X(\prod_{i=1}^{d}(-\infty,x_i])=\mathbb{P}(X_1\le x_1,\cdots,X_d\le x_d)$

Def. (同分布) 设 r.v.s $X,Y:(\Omega,\mathcal{F})\to(S,\mathcal{S})$. 若 $P_X=P_Y$ on $\mathcal{S}$,

则称 X 与 Y 同分布, 记 $X\overset{d}{=}Y$

Rmk. $P_X=\mathbb{P}(X^{-1}(\cdot))$, 若 X 在 $(\Omega,\mathcal{F},\mathbb{P})$ 上一个 r.v.

$P_Y=\mathbb{Q}(Y^{-1}(\cdot))$, 若 Y 在 $(\Omega,\mathcal{F},\mathbb{Q})$ 上一个 r.v.

Lem 设 r.v.s $X,Y:(\Omega,\mathcal{F})\to(\mathbb{R},\mathcal{B}_{\mathbb{R}})$ 在 $(\Omega,\mathcal{F},\mathbb{P})$ 上.

若 $F_X(x)=F_Y(x),\ \forall x\in\mathbb{R}$, 则 $X\overset{d}{=}Y$.

Def. (几乎处处相等). 设 r.v.s $X,Y:(\Omega,\mathcal{F})\to(S,\mathcal{S})$, 若 $\mathbb{P}(X=Y)=1$, 则

称 X 与 Y 几乎处处相等, 记 $X\overset{a.e.}{=}Y$.

Rmk. $X\overset{a.e.}{=}Y\Rightarrow X\overset{d}{=}Y$

• 设 μ 为 $(\mathbb{R},\mathcal{B}_{\mathbb{R}})$ 上概率测度, 定义 $F(x)=\mu((-\infty,x])$, $\forall x\in\mathbb{R}$,

则 $F(x)$ 为一个分布函数. (单增、右连续. $F(-\infty)=0$, $F(+\infty)=1$.)

$$\text{且}\begin{cases}\mu((a,b])=F(b)-F(a). \\ \mu((a,b))=F(b-)-F(a). \\ \mu([a,b])=F(b)-F(a-) \\ \mu([a,b))=F(b-)-F(a-)\end{cases}$$

这4个式子等价. $\forall -\infty<a<b<+\infty$.

反之, 通过上式之一建立概率测度 μ:

Eg. 设 r.v. X 取值于 $\{a_n : n\in\mathbb{Z}\}$ 且 $b_n = \mathbb{P}(X=a_n)$

则 X 的分布函数 $F_X(x) \triangleq \mathbb{P}(X\le x) = F_d(x)$ 为离散型分布函数.

定义 $F_c(x) \triangleq F(x) - F_d(x),\ x\in\mathbb{R}$，

Rmk. F_c 单增，$F_c(-\infty)=0$，$F_c(+\infty)=1-\sum b_n \in [0,1]$

$\Delta F_c(x) = \Delta F(x) - \Delta F_d(x) = 0,\ \forall x\in\mathbb{R} \Rightarrow F_c$ 连续

若 $F_c(+\infty)=1$，则 $F_c(x)$ 为一个连续型分布函数

首先，$\forall -\infty<a<b<+\infty$，$\mu([a,b)) \triangleq F(b-)-F(a)$.

step 1: 若 $C \overset{\text{开}}{\subset} \mathbb{R}$，则 $C \overset{\exists}{=} \bigcup_{i=1}^{\infty}(a_i,b_i)$

$$\mu(C) \triangleq \sum_{i=1}^{\infty}\mu((a_i,b_i)) = \sum_{i=1}^{\infty}[F(b_i-)-F(a_i)]$$

step 2: 若 $D \overset{\text{闭}}{\subset} \mathbb{R}$，则 D^c 开，$D^c \overset{\exists}{=} \bigcup_{i=1}^{\infty}(c_i,d_i)$

$$\mu(D) \triangleq 1-\mu(D^c) = 1-\sum_{i=1}^{\infty}[F(d_i-)-F(c_i)]$$

step 3: 若 $S\subset\mathbb{R}$，则

$$\mu^*(S) \triangleq \inf_{S\subset C\text{开}} \mu(C) \text{ ……外测度.}$$

$$\mu_*(S) \triangleq \sup_{S\supset D\text{闭}} \mu(D) \text{ ……内测度.}$$

若 $\mu_*(S)=\mu^*(S)$，称 S 关于分布函数 $F(x)$ "可测". （一般 $\mu_*(S)\le\mu^*(S)$）

当 S "可测"，$\mu(S) \triangleq \mu^*(S) = \mu_*(S)$.

Rmk. 1. 开集、闭集"可测"

2. $\mathcal{G} \triangleq \{S\subset\mathbb{R} : S \text{"可测"}\}$，则 $\mathcal{G}$ 为 σ-代数，从而 $\mathcal{B}_{\mathbb{R}} \subset \mathcal{G}$.

3. 上面定义 μ 为 $\mathcal{G}$ 上的概率测度.

§2.2 分布函数的性质：设 d.f. $F(x),\ x\in\mathbb{R}$.

Prop. (1) F 单增，从而 $F(x-) \triangleq \lim_{y\nearrow x} F(y) = \sup_{y<x} F(y)$

$$F(x+) \triangleq \lim_{y\searrow x} F(y) = \inf_{y>x} F(y)$$

$$F(x-) \le F(x) \overset{\text{(右连续)}}{=} F(x+)$$

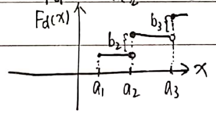
(2) 不连续点的类型（可去、跳跃、本质（左右极限不存在））为跳跃不连续点且可数.

Def. $\Delta F(x) \triangleq F(x+)-F(x-) = F(x)-F(x-) \ge 0$.

• 分布函数的分类与分解.

设 $F(x)$ 不连续点为 $\{a_n, n\in\mathbb{Z}\}$，$(\cdots<a_0<a_1<\cdots)$，$b_n \triangleq \Delta F(a_n) > 0$.

定义 $F_d(x) \triangleq \sum_{n\in\mathbb{Z}} b_n \mathbb{1}_{[a_n,+\infty)}(x),\ x\in\mathbb{R}$，其中 $\mathbb{1}_{[a_n,+\infty)}(x)$ 为一个退化的分布函数



Rmk. F_d 单增、右连续，$F_d(-\infty)=0$，$F_d(+\infty)=\sum b_n \in [0,1]$.

若 $F_d(+\infty)=1$，则称 F_d 为一个离散型分布函数

Pf (Jordan): 考虑 $\sum b_n \in (0,1)$ 定义 $\alpha = \sum b_n$，$F_1(x) = \frac{1}{\alpha} F_d(x)$，$F_2(x) = \frac{1}{1-\alpha} F_c(x)$ 即可 #

Pf of Lebesgue: 由 Jordan 分解，$F(x) = \alpha F_d(x) + (1-\alpha) F_c(x)$，其中 F_d 离散 d.f.

F_c 连续 d.f.

由单调 Lebesgue 分解，

$$F_c(x) = \hat{F}_{ac}(x) + \hat{F}_{cs}(x)，其中 \hat{F}_{ac}(x) \triangleq \int_{-\infty}^{x} F_c'(y)dy,$$

$$\hat{F}_{cs}(x) \triangleq F_2(x) - \hat{F}_{ac}(x).$$

则 $\hat{F}_{ac}(+\infty) = \int_{-\infty}^{+\infty} \hat{F}_{ac}'(y)dy \le F_c(+\infty) - F_c(-\infty) = 1$

$\hat{F}_{ac}$，$\hat{F}_{cs}$ 单增递减，$\hat{F}_{cs}' = 0$ a.e.，$\hat{F}_{ac}(-\infty) = \hat{F}_{cs}(-\infty) = 0$，$\hat{F}_{cs}(+\infty) \in [0,1]$

考虑 $\hat{F}_{ac}(+\infty) \in (0,1)$. 定义 $\beta = \hat{F}_{ac}(+\infty) \in (0,1)$. 取 $F_{ac} \triangleq \frac{1}{\beta}\hat{F}_{ac}$，$F_{cs} \triangleq \frac{1}{1-\beta}\hat{F}_{cs}$ 即可. #.

$\therefore F(x) = F_d(x) + F_c(x)$，$F_d(x) \triangleq \sum_{n\in\mathbb{Z}} b_n 1_{[a_n,+\infty)}(x)$，$F_c(x) \triangleq F(x) - F_d(x)$.

其中 F_d 单增、右连、$F_d(-\infty) = 0$，$F_d(+\infty) = \sum b_n \in [0,1]$

F_c … $F_c(-\infty) = 0$，$F_c(+\infty) = 1 - \sum b_n \in [0,1]$.

<u>Thm</u> (分布函数 Jordan 分解) 设分布函数 $F(x)$，$x \in \mathbb{R}$，则 $\exists\, \alpha \in (0,1)$，s.t.

$$F(x) = \alpha F_1(x) + (1-\alpha) F_2(x).$$

其中 F_1 为离散分布函数，F_2 为连续型分布函数.

<u>Def</u>. (AC-绝对连续型分布函数) 分布函数 $F(x)$. 若 $F(x) \in AC$. 即:

$\forall -\infty < x_1 < y_1 < \cdots < x_m < y_m < +\infty$，$m \ge 1$，$\varepsilon > 0$，$\exists\, \delta > 0$，s.t.

$$\sum_{i=1}^{m} |y_i - x_i| < \delta \Rightarrow \sum_{i=1}^{m} |F(y_i) - F(x_i)| < \varepsilon$$

则称 $F(x)$ 为一个 AC 型 d.f.

<u>Rmk</u>. $F \in$ A.C d.f $\Rightarrow \exists$ 非负 $f \in L^1(\mathbb{R})$，s.t. $\int_{x_1}^{x_2} f(x)dx = F(x_2) - F(x_1)$，$\forall -\infty \le x_1 < x_2 \le +\infty$

$\Rightarrow \quad \|f\|_{L^1(\mathbb{R})} \triangleq \int_{-\infty}^{+\infty} f(x)dx = F(+\infty) - F(-\infty) = 1$.

称 f 为 AC.d.f. $F(x)$ 的<u>密度函数</u>. $(F(dx) \ll dx)$

$\left(\frac{dF(x)}{dx} = f \ \cdots\cdots \text{Radon-Nikodym 导数}\right)$

<u>Def</u>. (奇异型 d.f.) 设 d.f. $F(x)$，若 $F' = 0$，则称 $F(x)$ 为奇异型分布函数.

若 $F(x)$ 还连续，则称 $F(x)$ 为连续 singular d.f.

<u>Rmk</u>. $\forall$ 离散 d.f. 都是 singular d.f.

连续 singular d.f.: Cantor d.f.

<u>Thm</u>. (单调函数的 Lebesgue 分解). 设 $f(x)$ 单调且 $f(-\infty) = 0$，f' 为 f 的导数(若存在)，

则 (i) $S \triangleq \{x \in \mathbb{R} : f'(x) \in [0,+\infty)\}$，$m(S^c) = 0$

(ii) $f' \in L^1(\mathbb{R})$，且 $\int_{x_1}^{x_2} f'(x)dx \le f(x_2) - f(x_1)$，"=" $\Leftrightarrow f \in AC$.

(iii) $f_{a.c.}(x) \triangleq \int_{-\infty}^{x} f'(y)dy$，$f_s(x) \triangleq f(x) - f_{a.c.}(x)$

则 $f_{a.c.} \in AC$，$f_{a.c.}$ 和 f_s 单增，$f_{a.c.}' = f'$ a.e.，$f_s' = 0$ a.e..

<u>Thm</u>. (d.f. 的 Lebesgue 分解.) 设 d.f $F(x)$，则 $\exists\, \alpha_1, \alpha_2 \in [0,1]$，s.t. $F(x) = \alpha_1 F_d(x) + \alpha_2 F_{ac}(x) + (1-\alpha_1-\alpha_2) F_{cs}(x)$

其中 F_d 离散 d.f.，F_{ac} AC d.f.，F_{cs} 连续奇异 d.f.

e.g. 设 $X \in L^1(\mathcal{F})$, $X \geq 0$ a.e. 定义 $\mu(A) \triangleq E(X \mathbb{1}_A) = \int_A X(\omega) \mathbb{P}(d\omega)$, $A \in \mathcal{F}$,

则 $\mu: \mathcal{F} \to [0,+\infty)$ 为一个有限测度.

i.e. ① $\mu(\phi) = 0$

② $\{A_n\} \subset \mathcal{F}$ disjoint $\Rightarrow E(X \mathbb{1}_{\cup A_n}) = \sum E(X \mathbb{1}_{A_n})$

③ $\mu(\Omega) = EX \in [0,+\infty)$.

当 $EX = 1$, μ 是 $\mathcal{F}$ 上概率测度

分布函数 $F(x)$ $\Leftrightarrow$ $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ 上概率测度 μ ($\mu = P_X$)

AC. d.f. $F(dx) \ll dx$ $\Rightarrow$ AC prob. measure $\mu \ll \nu$

singular d.f. $\Rightarrow$ singular prob. measure $\mu \perp \nu$

d.f. Lebesgue 分解 $\Rightarrow$ σ-有限 measure Lebesgue 分解 μ, ν.

$F(x) = F_{ac}(x) + F_{cs}(x)$ $\qquad$ $\nu = \nu_{ac} + \nu_s$, 其中 $\nu_{ac} \ll \mu$, $\nu_s \perp \mu$.

§2.3 积分(数学期望)

· Q: $X(\omega): (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ 为 $(\Omega, \mathcal{F}, P)$ 下的 r.v., 如何定义 X 关于 $\mathbb{P}$ 的积分

$$EX \triangleq \int_\Omega X \, d\mathbb{P} = \int_\Omega X(\omega) \mathbb{P}(d\omega) \ ?$$

· A: Step 1. 设 r.v. X 非负简单: $X(\omega) = \sum_{n \geq 1} b_n \mathbb{1}_{A_n}(\omega)$

定义 $EX = \int_\Omega X \, d\mathbb{P} = \int_\Omega X(\omega) \mathbb{P}(d\omega)$

$\triangleq \sum_{n \geq 1} b_n \mathbb{P}(A_n) \in [0,+\infty]$.

Step 2. 设 r.v. X 非负, 则 $\exists$ 非负简单 r.v.s $\{X_n\}_{n \geq 1}$, s.t. $X_n(\omega) \uparrow X(\omega)$, $\forall \omega \in \Omega$

定义. $EX \triangleq \begin{cases} \lim_{n\to\infty} EX_n, & \text{极限存在} \\ +\infty, & \text{否则}. \end{cases}$

Step 3. 设 r.v. $X = X^+ - X^-$, 定义 $EX \triangleq EX^+ - EX^-$

Rmk (1) 称 r.v. X 是可积的, 若 EX 存在.

可积 $\Leftrightarrow$ $L^p(\mathcal{F}) \triangleq \{X \in \mathcal{F} : E|X|^p < +\infty\}$, $p \geq 1$. $X \in L^1(\mathcal{F})$.

(2). $X \in Y \in L^1(\mathcal{F}) \Rightarrow \forall a, b \in \mathbb{R}$, $E(aX + bY) = aEX + bEY$

(3) $X \in L^1(\mathcal{F})$, $X \geq 0$ a.e. $(\mathbb{P}(X \geq 0) = 1.)$ $\Rightarrow EX \geq 0$

$X = 0$ a.e. $\Rightarrow EX = 0$.

$X, Y \in L^1(\mathcal{F})$, $X \geq Y$ a.e. $\Rightarrow EX \geq EY$

(4) X 可积 $\Leftrightarrow E|X| < +\infty$

(5) $X \in L^1(\mathcal{F})$, $A \in \mathcal{F}$ 且 $a \leq X(\omega) \leq b$, $\forall \omega \in A$. $\Rightarrow a\mathbb{P}(A) \leq \underset{\int_A X d\mathbb{P}}{\underbrace{EX\mathbb{1}_A}} \leq b\mathbb{P}(A)$

(6) $X \in L^1(\mathcal{F}) \Rightarrow |EX| \leq E|X|$.

Pf of Fatou: $\underline{\lim}_{n\to\infty} X_n(\omega) = \lim_{n\to\infty}\inf_{m\geq n} X_m(\omega)$

Let $Y_n = \inf_{m\geq n} X_m$, then $Y_n \uparrow \underline{\lim} X_n$ 且Y_n非负.

由MCT, $LHS = \lim E(Y_n) \leq \underline{\lim}\, E(X_n) = RHS$. #.

Pf of DCT: $-Y \leq X_n \leq Y$ a.e.

令 $Z_n = X_n + Y \geq 0$, $\hat{Z}_n = Y - X_n \geq 0$

则由Fatou, $E(\underline{\lim} Z_n) \leq \underline{\lim}\, EX_n + EY \Rightarrow E(\lim X_n) \leq \underline{\lim}\, X_n$

$E(\underline{\lim} \hat{Z}_n) \leq \underline{\lim}(-EX_n) + EY \Rightarrow E(\lim X_n) \geq \overline{\lim}\, X_n$ #.

Pf of $\mu \ll P$: 令 $X_n \triangleq X \wedge n = \min\{X, n\}$, 则 $X_n \uparrow X$ 且 $0 \leq X_n \leq n$.

$\therefore 0 \leq E[X_n \mathbb{1}_A] \leq n \cdot \mathbb{P}(A) = 0 \Rightarrow E[X_n \mathbb{1}_A] = 0$.

$\therefore E[X\mathbb{1}_A] = E[\lim X_n \mathbb{1}_A] \overset{MCT}{=} \lim E[X_n \mathbb{1}_A] = 0$.

Ex. 设$X \in L^1(\mathcal{F})$非负. 若$A_n \in \mathcal{F}$, $\mathbb{P}(A_n) \to 0$, 则$\lim E[X\mathbb{1}_{A_n}] = 0$.

Pf of $P \ll \mu$: $\mathbb{P}(A) = \mathbb{P}(A \cap \{X > 0\})$, 下证它为0.

令 $A_n = \{X > \frac{1}{n}\}$, $A_0 = \{X > 0\}$, 则 $A_n \uparrow A_0$

$\therefore 0 = E[X\mathbb{1}_A] \geq E[X \mathbb{1}_{A\cap A_n}] \geq \frac{1}{n} P(A \cap A_n) \geq 0 \Rightarrow \mathbb{P}(A\cap A_n) = 0$

$\therefore \mathbb{P}(A \cap A_0) = \mathbb{P}(A \cap (\bigcup_{n=1}^{\infty} A_n)) = \mathbb{P}(\bigcup_{n=1}^{\infty}(A\cap A_n)) = \lim_{n\to\infty} \mathbb{P}(A\cap A_n) = 0$.

e.g. 设$X \in L^1(\mathcal{F})$非负, 则 $EX = \int_\Omega X(\omega)\mathbb{P}(d\omega) = \int_\Omega \int_0^{X(\omega)} 1\,dt\, \mathbb{P}(d\omega)$

$= \int_\Omega \left(\int_0^{+\infty} \mathbb{1}_{\{X(\omega)\geq t\}} dt\right) \mathbb{P}(d\omega)$

$\overset{Fubini}{=} \int_0^{+\infty} \left(\int_\Omega \mathbb{1}_{\{X(\omega)\geq t\}} \mathbb{P}(d\omega)\right) dt$

$= \int_0^{+\infty} \mathbb{P}(X \geq t)\, dt$

Thm. (单调收敛 MCT). 设r.v.s $\{X_n\}_{n\geq 1}$ in $(\Omega, \mathcal{F}, \mathbb{P})$ 非负(可积)

且 $X_n \uparrow$ a.e., 则 $E[\lim X_n] = \lim EX_n$. (证明略)

Rmk. $X_0 \leq X_n \uparrow$ $\forall n \geq 1$, $X_0 \in L^1(\mathcal{F})$. 令 $Y_n \triangleq X_n - X_0$, 则MCT可去非负条件

Lem (Fatou) r.v.s $\{X_n\}_{n\geq 1}$ in $(\Omega, \mathcal{F}, P)$ 非负(可积), 则

$E(\underline{\lim} X_n) \leq \underline{\lim}\, EX_n$.

Thm. (控制收敛 DCT). r.v.s $\{X_n\}_{n\geq 1}$ in $(\Omega, \mathcal{F}, P)$ 中(可积), 满足

(i) $\forall n \geq 1$, $|X_n| \leq Y$, a.e. 其中$Y \in L^1(\mathcal{F})$非负.

(ii) $X_n \xrightarrow{a.e.} X$ 为r.v. $(P(\lim X_n = X) = 1)$

则 $E[\lim X_n] = \lim EX_n$

Rmk 有界收敛BCT: 取 $Y = M$ (常数).

Lem ($\mu \ll P$) 设 $X \in L^1(\mathcal{F})$ 非负r.v.. 若 $A \in \mathcal{F}$, $P(A) = 0$, 则 $E[X\mathbb{1}_A] = 0$.

Rmk. $\mu \ll P$: 称μ关于P绝对连续, i.e. $\forall A \in \mathcal{F}$, $P(A) = 0 \Rightarrow \mu(A) = 0$.

[AC函数 $F(x) = \int_{-\infty}^{x} f(y)dy$, $f \in L^1(\mathbb{R})$非负.

$\mu(A) \triangleq \int_A f(y)dy \Rightarrow \mu \ll P \Rightarrow \exists$非负Radon-Nikodym函数$\frac{d\mu}{dx} = f$]

Lem. ($P \ll \mu$) 设 $X \in L^1(\mathcal{F})$, $X > 0$ a.e. $A \in \mathcal{F}$. 若$E[X\mathbb{1}_A] = 0$, 则 $\mathbb{P}(A) = 0$.

Rmk. 此时, $\mu \sim P$: μ与P等价.

Cor.1 设$X \in L^1(\mathcal{F})$, $\forall A \in \mathcal{F}$. $E[X\mathbb{1}_A] = 0$, 则$X = 0$, a.e. (考虑$A_0 = \{X > 0\}$和上面lem.)

Cor 2. 设 X_1, X_2 为$(\Omega, \mathcal{F}, \mathbb{P})$中r.v.s, 且$\exists p > 0$, s.t. $E|X_1 - X_2|^p = 0$, 则$X_1 \overset{a.e.}{=} X_2$

(令 $A_0 = \{X_1 \neq X_2\}$, $X \triangleq |X_1 - X_2|^p > 0$ on A_0, 则 $E[X\mathbb{1}_{A_0}] = 0$, 由Lem. $\mathbb{P}(A_0) = 0$)

Cor 3. 设 X_1, X_2为$(\Omega, \mathcal{F}, \mathbb{P})$中r.v.s, 且$\forall A \in \mathcal{F}$, $E[X_1\mathbb{1}_A] \leq E[X_2 \mathbb{1}_A]$, 则$X_1 \overset{a.e.}{\leq} X_2$.

(令$A_0 = \{X_1 > X_2\}$, $X \triangleq X_1 - X_2 > 0$ on A_0, 则 $E[X\mathbb{1}_{A_0}] \leq 0$, $E[X\mathbb{1}_{A_0}] \geq 0$ $\Rightarrow E[X\mathbb{1}_{A_0}] = 0 \overset{Lem.}{\Downarrow} \mathbb{P}(A_0) = 0$.)

§2.4 变量变换公式.

Pf of 变量变换:

Step1 $h(x)=\mathbb{1}_B(x)$, $B\in\mathcal{S}$.

$$Eh(X)=E\mathbb{1}_B(X)=\mathbb{P}(X\in B)=P_X(B)=\int_S \mathbb{1}_B(x)P_X(dx)$$
$$=\int_S h(x)P_X(dx).$$

Step2. $h(x)$ 简单可测, 非负. $h(x)=\sum_{n=1}^{\infty} a_n \mathbb{1}_{B_n}(x)$.

$$Eh(X)=E\left[\sum_{n\ge1} a_n\mathbb{1}_{B_n}(X)\right]\overset{MCT}{=}\sum_{n=1}^{\infty}E[a_n\mathbb{1}_{B_n}(X)]$$
$$=\sum_{n=1}^{\infty}a_n\int_S \mathbb{1}_{B_n}(x)P_X(dx)$$
$$\overset{MCT}{=}\int_S\sum_{n\ge1}a_n\mathbb{1}_{B_n}(x)P_X(dx)=\int_S h(x)P_X(dx)$$

Step3. $h(x)$ 可测, 非负, $\exists$ 非负、简单可测 $h_n\uparrow h$, $\forall x\in\mathbb{R}$.

$$Eh(X)\overset{MCT}{=}\lim_{n\to\infty}Eh_n(X)\overset{MCT}{=}\int_S h(x)P_X(dx).$$

Step4. $h(x)=h^+(x)-h^-(x)$

$$Eh^+(X)=\int_S h^+(x)P_X(dx),\ Eh^-(X)=\int_S h^-(x)P_X(dx)$$
$$\Rightarrow E[h(X)]=E[h^+(X)]-E[h^-(X)]=\int_S h(x)P_X(dx) \qquad \#$$

E.g ① r.v. X 连续型, i.e., $\exists$ 非负 $f\in L^1(\mathbb{R})$ 且 $\|f\|_{L^1(\mathbb{R})}=1$, s.t. $F_X(x)=\int_{-\infty}^{x}f(y)dy$ (AC d.f.).

则 X 的分布 $P_X(B)=\int_B f(y)dy$, $B\in\mathcal{B}_{\mathbb{R}}$

$$P_X(dx)=F_X(dx)=f(x)dx.$$

变量变换: $E[h(X)]=\int_{\mathbb{R}}h(x)P_X(dx)=\int_{\mathbb{R}}h(x)f(x)dx.$

② r.v. X 离散型, 取值 $\{a_n\}$, 且分布 $b_n=\mathbb{P}(X=a_n)$, $\sum_{n\in\mathbb{Z}}b_n=1$

则 X 的分布函数 $F_X(x)=\sum_{n\in\mathbb{Z}}b_n\mathbb{1}_{[a_n,+\infty)}(x)$

$$P_X=\sum b_n\delta_{a_n},\ \text{其中 } \delta_{a_n} \text{ 为 Dirac-delta 测度}$$

i.e., $P_X(B)=\sum b_n\delta_{a_n}(B)=\sum b_n\mathbb{1}_B(a_n)$

变量变换: $E[h(X)]=\int_{\mathbb{R}}h(x)\sum b_n\delta_{a_n}(dx)=\sum b_n h(a_n)$

Thm (变量变换公式) 设 r.v. $X:(\Omega,\mathcal{F})\to(S,\mathcal{S})$, 可测 $h:(S,\mathcal{S})\to(\mathbb{R},\mathcal{B}_{\mathbb{R}})$

且 $h(X)\in L^1(\mathcal{F})$, 则

$$E[h(X)]=\int_\Omega h(X)(\omega)P(d\omega)=\int_S h(x)P_X(dx)$$

其中 P_X 为 r.v. X 的分布

特别, $(S,\mathcal{S})=(\mathbb{R},\mathcal{B}_{\mathbb{R}})$ 时, $E[h(X)]=\int_{\mathbb{R}}h(x)F_X(dx)$,

其中 $F_X(x)\triangleq P_X((-\infty,x])$, $x\in\mathbb{R}$ 为 r.v. X 的 d.f.

Rmk. Def. (Push forward measure.) 设 $T:(S_1,\mathcal{S}_1)\to(S_2,\mathcal{S}_2)$ 可测映射.

μ 为 $(S_1,\mathcal{S}_1)$ 上的一个测度. 称

$(T\#\mu)(B)\triangleq\mu(T^{-1}(B))$, $\forall B\in\mathcal{S}_2$ 为测度 μ 关于 T 的 pfm.

$\therefore P_X$ 为 $\mathbb{P}$ 关于 X 的 pfm. ($P_X(B)=\mathbb{P}(X^{-1}(B))$, $\forall B\in\mathcal{S}$)

Thm (General-变量变换) 设 h 在 $(S_2,\mathcal{S}_2)$ 上可测, 则 h 关于 $T\#\mu$ 可积

$\Leftrightarrow$ $h\circ T$ 关于 μ 可积

且有 $\int_{S_2}h\,d(T\#\mu)=\int_{S_1}h\circ T\,d\mu$

Rmk. 若 $(S_i,\mathcal{S}_i)=(S,\mathcal{S})$, $i=1,2$, $T\#\mu=\mu$.

则称 μ 关于 T 是不变的 (或 μ 为 T 的不变测度.)

e.g. $(S,\mathcal{S})=(\mathbb{R},\mathcal{B}_{\mathbb{R}})$, $T^a(x)=x+a$ (translation map), m 为 $\mathcal{B}_{\mathbb{R}}$ 上测度

则 $(T^a_\# m)(B)=m((T^a)^{-1}(B))=m(\{x\in\mathbb{R}: x+a\in B\})$

$=m(B-a)=m(B).$

Lebesgue 测度 m 为 T^a 的一个不变测度.

§2.5 独立性与乘积测度.

Def (σ-代数独立性) 设 $\mathcal{H},\mathcal{G}\subset\mathcal{F}$ 为 σ-代数, 若 $\forall H\in\mathcal{H}, G\in\mathcal{G}$

$$P(H\cap G)=P(H)P(G)$$

则称 $\mathcal{H}$ 和 $\mathcal{G}$ 相互独立.

（σ-代数独立性与事件独立性等价）

e.g. 设 $X=\mathbb{1}_A$，$Y=\mathbb{1}_B$，$A,B\in\mathcal{F}$，则

$\sigma(X)=\{\phi,\Omega,A,A^c\}$

$\sigma(Y)=\{\phi,\Omega,B,B^c\}$

A 与 B 独立 $\Leftrightarrow \sigma(X)$ 与 $\sigma(Y)$ 独立.

Pf: $\forall L\le n-1$，$i_1,\cdots,i_L\subset\{1,\cdots,n-1\}$，$H\triangleq\bigcap_{k=1}^{L}A_{i_k}$，$A_{i_k}\in\mathcal{A}_{i_k}$

$\forall G\in\mathcal{G}_n=\sigma(\mathcal{A}_n)$

$$\begin{cases}\mu_1(G)\triangleq\mathbb{P}(H\cap G)\\ \mu_2(G)\triangleq\mathbb{P}(H)\cdot\mathbb{P}(G)\end{cases}$$

则 μ_1,μ_2 为 $\mathcal{G}_n$ 上两个有限测度，$\mu_1(\Omega)=\mu_2(\Omega)=\mathbb{P}(H)$

由独立性 $\mu_1=\mu_2$ on $\mathcal{A}_n$.

令 $\mathcal{L}=\{G\in\mathcal{G}_n:\mu_1(G)=\mu_2(G)\}$，则 $\mathcal{L}$ 为 λ 类，$\mathcal{A}_n\subset\mathcal{L}$.

由 π-λ 定理，$\mu_1=\mu_2$ on $\mathcal{G}_n$.

$\therefore\forall G\in\mathcal{G}_n$，$\mathbb{P}(H\cap G)=\mathbb{P}(H)\cdot\mathbb{P}(G)\Rightarrow\mathcal{A}_1,\cdots,\mathcal{A}_{n-1},\mathcal{G}_n$ 独立.

$\mathcal{G}_n,\mathcal{A}_1,\cdots,\mathcal{A}_{n-2},\mathcal{A}_{n-1}$

$\Rightarrow$ 归纳，$\mathcal{G}_1,\cdots,\mathcal{G}_n$ 相互独立 #.

Ex. $X_1,\cdots,X_n$ 为 $(\Omega,\mathcal{F},\mathbb{P})$ 下的 r.v.s，则 $X_1\cdots X_n$ 相互独立 $\Leftrightarrow$

$$\mathbb{P}(X_1\le x_1,\cdots,X_n\le x_n)=\prod_{k=1}^{n}\mathbb{P}(X_k\le x_k),\ \forall x_k\in\mathbb{R},\ k=1,\cdots,n.$$

Pf: $X_1\cdots X_n$ 相互独立 $\Leftrightarrow\sigma(X_1)\cdots\sigma(X_n)$ 相互独立

$(\Rightarrow)$: $B_k=(-\infty,x_k]\in\mathcal{B}_{\mathbb{R}}$，$\mathbb{P}[\bigcap X_k^{-1}((-\infty,x_k])]=\prod_{k=1}^{n}\mathbb{P}(X_k^{-1}((-\infty,x_k]))$

$(\Leftarrow)$: $\mathcal{A}_k\triangleq\{X_k^{-1}((-\infty,x]),x\in\mathbb{R}\}$，则 $\mathcal{A}_k$ 为 π-类

$\because\mathbb{P}(\bigcap_{k=1}^{n}X_k^{-1}((-\infty,x_k]))=\prod_{k=1}^{n}\mathbb{P}(X_k^{-1}((-\infty,x_k]))\Rightarrow$ 可把 $k=1,\cdots,n$ 换成 $k=i_1,\cdots,i_L$，$L\le n$

$\therefore\mathcal{A}_1\cdots\mathcal{A}_n$ 相互独立.

$\because\sigma(X_k)=\sigma(\mathcal{A}_k)$ $\left(\sigma(X_k)=\{X_k^{-1}(\sigma(\{(-\infty,x_k]\}))\}=\sigma\{X_k^{-1}(\{(-\infty,x_k]\})\}\right)$

$\therefore$ 由右边 Thm，$\sigma(X_1)\cdots\sigma(X_n)$ 相互独立. #

Def（r.v. 独立性）设 r.v.s X,Y on $(\Omega,\mathcal{F},\mathbb{P})$. 若 $\sigma(X)$ 与 $\sigma(Y)$ 独立，

则称 r.v.s X 与 Y 相互独立.

Def.（多个事件独立性）设 $\{A_n\}\subset\mathcal{F}$，若 $\forall L\ge1$，$A_{i_1},\cdots,A_{i_L}$ 互不相同，

$$\mathbb{P}(\bigcap_{k=1}^{L}A_{i_k})=\prod_{k=1}^{L}\mathbb{P}(A_{i_k}),$$

则称 $\{A_n\}$ 为相互独立的.

Rem $\{A_n\}$ 可以是 $n\in\mathbb{N}$ 或 $n\in I$.

Thm. 设 $\mathcal{G}_i\triangleq\sigma(\mathcal{A}_i)\subset\mathcal{F}$，$i=1,\cdots,n$. 若 $\mathcal{A}_1,\cdots\mathcal{A}_n$ 相互独立的 π-类，

则 $\mathcal{G}_1,\cdots,\mathcal{G}_n$ 相互独立.

Cor. 设 $\mathcal{G}_\alpha=\sigma(\mathcal{A}_\alpha)\subset\mathcal{F}$，$\alpha\in I$. 若 $\{A_\alpha\}_{\alpha\in I}$ 相互独立的 π 类，

则 $\{\mathcal{G}_\alpha\}_{\alpha\in I}$ 相互独立.

Ex r.v.s $X_1,\cdots,X_n:(\Omega,\mathcal{F},\mathbb{P})\to(S,\mathcal{S})$ 独立 且 $f_1,\cdots,f_n:(S,\mathcal{S})\to(T,\mathcal{T})$ 可测.

则 $f_1(X_1)\cdots f_n(X_n)$ 相互独立.

Pf: 由于 $f_1(X_1)\cdots f_n(X_n):(\Omega,\mathcal{F})\to(T,\mathcal{T})$ 为 r.v.s，只须证它们相互独立

$\Leftrightarrow\sigma(f_1(X_1))\cdots\sigma(f_n(X_n))$ 相互独立.

$\because\forall C_k\in\mathcal{T}$，$\mathbb{P}(\bigcap_{k=1}^{n}(f_k\circ X_k)^{-1}(C_k))=\mathbb{P}(\bigcap_{k=1}^{n}X_k^{-1}[f_k^{-1}(C_k)])$

$$\overset{f_k^{-1}(C_k)\in\mathcal{S}}{\underset{X_k\text{独立}}{=}}\prod_{k=1}^{n}\mathbb{P}(X_k^{-1}[f_k^{-1}(C_k)])=\prod_{k=1}^{n}\mathbb{P}((f_k\circ X_k)^{-1}(C_k))$$ #

Thm. $i\ge1$，$m(i)\ge1$，$\{\mathcal{F}_{ij}\}_{i\ge1,j=1,\cdots,m(i)}$ 为相互独立的 σ-代数.

$\mathcal{F}_i\triangleq\bigvee_{j=1}^{m(i)}\mathcal{F}_{ij}\left(=\sigma(\bigcup_{j=1}^{m(i)}\mathcal{F}_{ij})\right)$，则 $\{\mathcal{F}_i\}_{i\ge1}$ 相互独立.

$\mathcal{F}_{11},\mathcal{F}_{12},\cdots,\mathcal{F}_{1,m(1)}\Rightarrow\mathcal{F}_1$

$\mathcal{F}_{21},\mathcal{F}_{22},\cdots,\mathcal{F}_{2,m(2)}\Rightarrow\mathcal{F}_2$

$\cdots\quad\cdots\quad\cdots$

$\mathcal{F}_{n1}\ \mathcal{F}_{n2}\ \cdots\ \mathcal{F}_{n,m(n)}\Rightarrow\mathcal{F}_n$

Pf: $\mathcal{Q}_i\triangleq\{\bigcap_{j=1}^{m(i)}E_{ij}:E_{ij}\in\mathcal{F}_{ij}\}$，则 $\{\mathcal{Q}_i\}$ 为相互独立的 π-类且 $\sigma(\mathcal{Q}_i)=\mathcal{F}_i$

$\therefore\{\mathcal{F}_i\}$ 相互独立. #.

Pf: (i) $\forall B\in\mathcal{S}^{\otimes n}$, $C\in\mathcal{S}$, $\mathbb{P}((X_1\cdots X_n)^{-1}(B)\cap X_{n+1}^{-1}(C))=\mathbb{P}((X_1\cdots X_n)^{-1}(B))\mathbb{P}(X_{n+1}^{-1}(C))$

取 $B=\prod_{k=1}^{n}B_k$, $B_k\in\mathcal{S}$, 则有 $\mathbb{P}(\bigcap_{k=1}^{n}X_k^{-1}(B_k)\cap X_{n+1}^{-1}(C))=\mathbb{P}(\bigcap_{k=1}^{n}X_k^{-1}(B_k))\mathbb{P}(X_{n+1}^{-1}(C))$

$\overset{\text{归纳}}{=}\mathbb{P}(X_1^{-1}(B_1))\cdots\mathbb{P}(X_{n+1}^{-1}(C))$

$\therefore \sigma(X_1)\cdots\sigma(X_{n+1})$ 相互独立, $\forall n$

$\therefore \sigma(X_1)\cdots\sigma(X_n)\cdots$ 相互独立.

(ii) 注意 $\mathcal{T}_n^X=\sigma(\bigcup_{d=1}^{\infty}\sigma(X_{n+1}\cdots X_{n+d}))$.

由 $X_1\cdots X_n\cdots$ 相互独立, $\sigma(X_1\cdots X_n)$ 与 $\sigma(X_{n+1}\cdots X_{n+d})$ 相互独立.

$\Rightarrow \sigma(X_1\cdots X_n)$ 与 $\bigcup_{d=1}^{\infty}\sigma(X_{n+1}\cdots X_{n+d})$ 相互独立.

$\because \bigcup_{d=1}^{\infty}\sigma(X_{n+1}\cdots X_{n+d})$ 是个π类 (由它是过滤之并)

$\therefore \sigma(X_1\cdots X_n)$ 与 $\sigma(\bigcup_{d=1}^{\infty}\sigma(X_{n+1}\cdots X_{n+d}))=\mathcal{T}_n^X$ 相互独立. #

Pf of Kolmogorov: 只须证 $\mathcal{T}^X$ 与 $\mathcal{T}^X$ 相互独立 ($\Rightarrow \forall H\in\mathcal{T}^X$, $P(H\cap H)=P(H)^2$ $\Rightarrow P(H)=0$ 或 1.)

已有: $\sigma(X_1\cdots X_n)$ 与 $\mathcal{T}_n^X$ 独立

$\Rightarrow \sigma(X_1\cdots X_n)$ 与 $\bigcap_{k=1}^{\infty}\mathcal{T}_k^X$ 独立

$\Rightarrow \bigcup_{n=1}^{\infty}\sigma(X_1\cdots X_n)$ 与 $\bigcap_{k=1}^{\infty}\mathcal{T}_k^X$ 独立.

$\overset{\text{LHS是过滤π类}}{\Rightarrow} \sigma(\bigcup_{n=1}^{\infty}\sigma(X_1\cdots X_n))$ 与 $\mathcal{T}^X$ 独立

$\Rightarrow \sigma(X_1\cdots X_n\cdots)$ 与 $\mathcal{T}^X$ 独立

$\overset{\text{RHS}\subset\text{LHS}}{\Rightarrow} \mathcal{T}^X$ 与 $\mathcal{T}^X$ 独立 #

Pf of section prop: 令 $\mathcal{M}\triangleq\{E\subset\Omega_1\times\Omega_1: E_{w_1}\in\mathcal{F}_2, E_{w_2}\in\mathcal{F}_1\}$ 为σ-代数

$\forall E=A\times B\in\mathcal{R}$, $E_{w_1}=\begin{cases}B, & w_1\in A\\ \phi, & w_1\notin A.\end{cases}\in\mathcal{F}_2$ $\quad\therefore \mathcal{R}\subset\mathcal{M}$

$E_{w_2}=\begin{cases}A, & w_2\in B\\ \phi, & w_2\notin B\end{cases}\in\mathcal{F}_1$ $\quad\therefore \sigma(\mathcal{R})\subset\mathcal{M}$ #

Pf of $\mathcal{B}_{\mathbb{R}^m}\otimes\mathcal{B}_{\mathbb{R}^n}=\mathcal{B}_{\mathbb{R}^{m+n}}$: Step 1: $\prod_{i=1}^{m+n}[a_i,b_i]=\prod_{i=1}^{m}[a_i,b_i]\times\prod_{i=m+1}^{m+n}[a_i,b_i]\in$ LHS. $\Rightarrow$ RHS $\subset$ LHS

Step 2: $\{A\times B: A\in\mathcal{B}_{\mathbb{R}^m}, B\in\mathcal{B}_{\mathbb{R}^n}\}=\{A\times\mathbb{R}^n; A\in\mathcal{B}_{\mathbb{R}^m}\}\cap\{\mathbb{R}^m\times B; B\in\mathcal{B}_{\mathbb{R}^n}\}$

下证 $\{A\times\mathbb{R}^n: A\in\mathcal{B}_{\mathbb{R}^m}\}\subset\mathcal{B}_{\mathbb{R}^{m+n}}$.

令 $\mathcal{M}\triangleq\{A\subset\mathbb{R}^m: A\times\mathbb{R}^n\in\mathcal{B}_{\mathbb{R}^{m+n}}\}$ 为σ-代数 且 $\{\mathbb{R}^m$中开集$\}\subset\mathcal{M}\Rightarrow\mathcal{B}_{\mathbb{R}^m}\subset\mathcal{M}$. #

Thm. 设 r.v.s $X_1\cdots X_n\cdots$ 为 $(\Omega,\mathcal{F},\mathbb{P})$ 下的, 则

(i) 若 $\forall n\geq 1$, $\sigma(X_1\cdots X_n)$ 与 $\sigma(X_{n+1})$ 独立, 则 $X_1\cdots X_n\cdots$ 相互独立

(ii) 若 $X_1\cdots X_n\cdots$ 相互独立, 则 $\sigma(X_1\cdots X_n)$ 与 $\mathcal{T}_n^X\triangleq\sigma(X_i, i>n)$ 相互独立.

Thm (Kolmogorov 0-1 律): $X_1\cdots X_n\cdots$ 为 $(\Omega,\mathcal{F},\mathbb{P})$ 下相互独立 r.v.s. 定义

$\mathcal{T}^X\triangleq\bigcap_{n=1}^{\infty}\mathcal{T}_n^X=\bigcap_{n=1}^{\infty}\sigma(X_i, i>n)$ 尾σ-代数.

则 $\mathcal{T}^X$ 为一平凡σ-代数, i.e. $\forall H\in\mathcal{T}^X$, $P(H)=0$ 或 1.

Rmk 事件形式的 Kolmogorov 0-1 律: 设 $A_1\cdots A_n\cdots$ 相互独立事件.

则 $\mathbb{P}(\overline{\lim}A_n)=0$ 或 1

(Hint: 定义 $X_n\triangleq\mathbb{1}_{A_n}$, 则 $X_1\cdots X_n\cdots$ 相互独立

定义 $\mathcal{T}^X\triangleq\bigcap_{n=1}^{\infty}\sigma(X_i, i>n)$, 由 $\overline{\lim}A_n\in\mathcal{T}^X$ 得证.)

乘积测度

Def. (可测矩形) 可测空间 $(\Omega_i,\mathcal{F}_i)_{i=1,2}$. 称 $A\times B$, $A\in\mathcal{F}_1$, $B\in\mathcal{F}_2$ 为可测矩形

记 $\mathcal{R}\triangleq\{A\times B: A\in\mathcal{F}_1, B\in\mathcal{F}_2\}$

Rmk ① $\forall A\times B, C\times D\in\mathcal{R}$, $(A\times B)\cap(C\times D)=(A\cap C)\times(B\cap D)$

$\Rightarrow\mathcal{R}$ 是π类.

② $\forall A\times B\in\mathcal{R}$, $(A\times B)^c=(A^c\times B)\cup(A\times B^c)\cup(A^c\times B^c)$

Def. (乘积σ-代数) 可测空间 $(\Omega_i,\mathcal{F}_i)_{i=1,2}$. 称 $\sigma(\mathcal{R})$ 为 $\mathcal{F}_1$ 与 $\mathcal{F}_2$ 的乘积σ-代数

记为 $\mathcal{F}_1\otimes\mathcal{F}_2\triangleq\sigma(\mathcal{R})$.

Rmk $(\Omega_1\times\Omega_2,\mathcal{F}_1\otimes\mathcal{F}_2)$ 为可测空间, 称 $\forall E\in\mathcal{F}_1\otimes\mathcal{F}_2$ $(\mathcal{F}_1\otimes\mathcal{F}_2)$ 可测.

特别, 任何可测矩形都可测.

Def. (截口). $E\subset\Omega_1\times\Omega_2$. 称 $E_{w_1}\triangleq\{w\in\Omega_2:(w_1,w_2)\in E\}\subset\Omega_2$, $\forall w_1\in\Omega_1$

为 E 的 w_1-截口 (w_1-section)

称 $E_{w_2}\triangleq\{w\in\Omega_1:(w_1,w_2)\in E\}\subset\Omega_1$, $\forall w_2\in\Omega_2$ 为 E 的 w_2-截口.

Lem (section prop.) 若 $E\in\mathcal{F}_1\otimes\mathcal{F}_2$, 则 $E_{w_1}\in\mathcal{F}_2$, $E_{w_2}\in\mathcal{F}_1$.

Lem $\mathcal{B}_{\mathbb{R}^m}\otimes\mathcal{B}_{\mathbb{R}^n}=\mathcal{B}_{\mathbb{R}^{m+n}}$

$=\sigma(\{A\times B: A\in\mathcal{B}_{\mathbb{R}^m}, B\in\mathcal{B}_{\mathbb{R}^n}\})$ $\quad\mathcal{B}_{\mathbb{R}^{m+n}}=\sigma(\{\mathbb{R}^{m+n}$中开(闭)集$\})$.

Pf of Lem in Step 2:

$\because \mathbb{1}_{A\times B}(w_1,w_2)=\mathbb{1}_{\coprod_{k=1}^{\infty} A_k\times B_k}(w_1,w_2)=\sum_{k=1}^{\infty}\mathbb{1}_{A_k\times B_k}(w_1,w_2)$

$\therefore \mathbb{1}_A(w_1)\mathbb{1}_B(w_2)=\sum \mathbb{1}_{A_k}(w_1)\mathbb{1}_{B_k}(w_2)$

由MCT,

$\mathbb{1}_A(w_1)\nu_2(B)=\mathbb{1}_A(w_1)\int_{\Omega_2}\mathbb{1}_B(w_2)\nu_2(dw_2)=\int_{\Omega_2}\sum \mathbb{1}_{A_k}(w_1)\mathbb{1}_{B_k}(w_2)\nu_2(dw_2)$

$\overset{MCT}{=}\sum\int_{\Omega_2}\mathbb{1}_{A_k}(w_1)\mathbb{1}_{B_k}(w_2)\nu_2(dw_2)$

$=\sum \mathbb{1}_{A_k}(w_2)\nu_2(B_k)$

同理,再由MCT, $\nu_1(A)\nu_2(B)=\sum \nu_1(A_k)\nu_2(B_k)$ #

Pf of 独立性: 不妨 $n=2$.

X_1与X_2独立 $\Leftrightarrow \sigma(X_1)$, $\sigma(X_2)$独立.

$\Leftrightarrow \forall B_i\in\mathcal{S}_i$, $\mathbb{P}(X_1^{-1}(B_1)\cap X_2^{-1}(B_2))=\mathbb{P}(X_1^{-1}(B_1))\mathbb{P}(X_2^{-1}(B_2))$

$\because LHS=\mathbb{P}((X_1,X_2)^{-1}(B_1,B_2))=P_{(X_1,X_2)}(B_1\times B_2)$

$RHS=P_{X_1}(B_1)P_{X_2}(B_2)=(P_{X_1}\times P_{X_2})(B_1\times B_2)$

$\therefore P_{(X_1,X_2)}=P_{X_1}\times P_{X_2}$ on R (π-类)

$\overset{\pi-\lambda定理}{\Rightarrow} P_{(X_1,X_2)}=P_{X_1}\times P_{X_2}$ on $\sigma(R)=\mathcal{S}_1\otimes\mathcal{S}_2$ #.

Rmk. of Fubini: 设 $X:(\Omega,\mathcal{F})\to(S_1,\mathcal{S}_1)$ 独立 r.v.s. $\Rightarrow P_{(X_1,X_2)}=P_{X_1}\times P_{X_2}$.
$Y:(\Omega,\mathcal{F})\to(S_2,\mathcal{S}_2)$ on $\mathcal{S}_1\otimes\mathcal{S}_2$

设 $h:S_1\times S_2\to\mathbb{R}$ 为 $\mathcal{S}_1\otimes\mathcal{S}_2$ 可测 且 ①非负/② $E|h(X,Y)|<\infty$.

则 $E[h(X,Y)]=\int_{S_1}\int_{S_2}h(x,y)P_{X_2}(dy)P_{X_1}(dx)$

$=\int_{S_2}\int_{S_1}h(x,y)P_{X_1}(dx)P_{X_2}(dy)$

特别, 若 $h(X,Y)=f(X)g(Y)$, 则 $E[f(X)g(Y)]=Ef(X)\cdot Eg(Y)$.
#.

下面定义乘积测度. ν_i 为可测空间 $(\Omega_i,\mathcal{F}_i)$ 上 σ-有限测度 $(i=1,2)$

Step 1. $\forall E=A\times B\in R$, $\nu(E)\triangleq\nu_1(A)\times\nu_2(B)$

Step 2. $\mathcal{A}\triangleq\{\coprod_{k=1}^{m}A_k\times B_k,\ A_k\in\mathcal{F}_1,\ B_k\in\mathcal{F}_2\}$, 则 $\mathcal{A}$ 为代数.

Lem $A\times B\in R$ 且 $A\times B=\coprod_{k=1}^{\infty}A_k\times B_k$, $A_k\in\mathcal{F}_1$, $B_k\in\mathcal{F}_2$.

则 $\nu(A\times B)=\sum_{k=1}^{\infty}\nu(A_k\times B_k)\overset{Step 1}{=}\sum_{k=1}^{\infty}\nu_1(A_k)\times\nu_2(B_k)$.

故可将 ν 从 R 延拓到 $\mathcal{A}$ 上. 由Lem及其证明方法, ν 在 $\mathcal{A}$ 上可列加.

Thm $(\Omega_i,\mathcal{F}_i,\underset{(\sigma-有限)}{\nu_i})$ $i=1,2$. $\exists 1$ $(\Omega_1\times\Omega_2,\mathcal{F}_1\otimes\mathcal{F}_2)$ 中 σ-有限测度

ν, s.t. $\nu(\coprod_{k=1}^{\infty}A_k\times B_k)=\sum_{k=1}^{\infty}\nu_1(A_k)\times\nu_2(B_k)$, $A_k\in\mathcal{F}_1$, $B_k\in\mathcal{F}_2$

(Hint: 由 Caratheodory, $\sigma(R)=\sigma(\mathcal{A})=\mathcal{F}_1\otimes\mathcal{F}_2$.)

· 记 $\nu_1\times\nu_2\triangleq\nu$ on $\mathcal{F}_1\otimes\mathcal{F}_2$

Thm. (独立性) $X_i:(\Omega,\mathcal{F})\to(S_i,\mathcal{S}_i)$ 为 $(\Omega,\mathcal{F},\mathbb{P})$ 下的 r.v.s, 则

$X_1\cdots X_n$ 相互独立 $\Leftrightarrow P_{(X_1\cdots X_n)}=P_{X_1}\times\cdots\times P_{X_n}$ on $\mathcal{S}_1\otimes\cdots\otimes\mathcal{S}_n$

Kolmogorov 延拓定理.

· 定义 $\mathbb{R}^{\infty}=\{x=(x_1,x_2,\cdots),x_i\in\mathbb{R}\}$

$R\triangleq\{\{x=(x_1,\cdots,x_n,\cdots)\in\mathbb{R}^{\infty}: x_i\in B_i\in\mathcal{B}_{\mathbb{R}}, i=1,\cdots n\}: n\geq 1\}$.

$\mathcal{B}_C\triangleq\sigma(R)$

则 $(\mathbb{R}^{\infty},\mathcal{B}_C)$ 为一个可测空间.

Thm (Kolmogorov) $\forall n\geq 1$, μ_n 为 $(\mathbb{R}^n,\mathcal{B}_{\mathbb{R}^n})$ 上 概率测度, 且

$\mu_{n+1}(B_1\times\cdots\times B_n\times\mathbb{R})=\mu_n(B_1\times\cdots\times B_n)$ 相容性条件. $B_i\in\mathcal{B}_{\mathbb{R}}$

则 $\exists 1$ $(\mathbb{R}^{\infty},\mathcal{B}_C)$ 上概率测度 μ_{∞}, s.t. $\mu_{\infty}(\{x=(x_1,\cdots,x_n,\cdots)\in\mathbb{R}^{\infty}: x_i\in B_i, i=1,\cdots,n\})$

$=\mu_n(B_1\times\cdots\times B_n)$

Thm (Fubini) 设 ν_i 为 $(\Omega_i,\mathcal{F}_i)$ 上 σ-有限测度, $i=1,2$. $\mu\triangleq\nu_1\times\nu_2$.

设 $h:\Omega_1\times\Omega_2\to\mathbb{R}$ 为 $\mathcal{F}_1\otimes\mathcal{F}_2$ 可测 且 ①非负/② $\int_{\Omega_1\times\Omega_2}|h|d\mu<\infty$.

则 $\int_{\Omega_1\times\Omega_2}h(w_1,w_2)\mu(dw_1,dw_2)=\int_{\Omega_1}\int_{\Omega_2}h(w_1,w_2)\nu_2(dw_2)\nu_1(dw_1)$

$=\int_{\Omega_2}\int_{\Omega_1}h(w_1,w_2)\nu_1(dw_1)\nu_2(dw_2)$.

引例. 设 (X,Z) 为 $(\Omega,\mathcal{F},\mathbb{P})$ 下取值于 $\{(a_i,b_j)\}_{j=1,\cdots,n}^{i=1,\cdots,m}$ 的离散型 r.v.

定义分布律 $p_{ij}=P(X=a_i,Z=b_j)$

则 $E[X\mid Z=b_j]=\sum_{i=1}^{n}a_i P(X=a_i\mid Z=b_j)=\dfrac{\sum_{i=1}^{m}a_i p_{ij}}{\sum_{i=1}^{m}p_{ij}}$ 是关于 b_j 的函数

记 $E[X\mid\sigma(Z)]$ 或 $E[X\mid Z]\triangleq f(Z)$ 为一 r.v.

注意到 $Z=\sum_{j=1}^{n}b_j\mathbb{1}_{G_j}$, $G_j=\{Z=b_j\}$

$\{G_j\}$ 为 Ω 的一个划分. $\Rightarrow Z$ 是一个简单 r.v.

则 $\sigma(Z)=\sigma(\{G_1,\cdots,G_n\})=\{\bigcup_{j\in J}G_j,\ J\subset\{1,\cdots,n\}\}$ (J 可取 ϕ)

$\forall G\in\sigma(Z)$, $\exists J$, s.t. $G=\bigcup_{j\in J}G_j$

$E[f(Z)\mathbb{1}_G]=\sum_{j\in J}E[f(Z)\mathbb{1}_{G_j}]=\sum_{j\in J}f(b_j)P(Z=b_j)$

$=\sum_{j\in J}\dfrac{\sum_{i=1}^{m}a_i p_{ij}}{P(Z=b_j)}P(Z=b_j)=\sum_{j\in J}\sum_i a_i p_{ij}$

$E[X\mathbb{1}_G]$

$=\sum_{j\in J}E[X\mathbb{1}_{G_j}]=\sum_{j\in J}\sum_{i=1}^{m}a_i p_{ij}$

Summary: $f(Z)=E[X\mid\sigma(Z)]$ 满足:

(i) $E[X\mid\sigma(Z)]$ 是 $\sigma(Z)$-可测的 ($\forall B\in\mathcal{B}_{\mathbb{R}}$, $f^{-1}(B)\in\sigma(Z)$.)

(ii) $\forall G\in\sigma(Z)$, $E\{E[X\mid\sigma(Z)]\mathbb{1}_G\}=E[X\mathbb{1}_G]$.

Pf of Lebesgue 分解: 不妨 μ,ν 为有限测度.

为使 ν_s 为测度 $\Rightarrow\forall B\in\mathcal{S}$, $\nu_s(B)=\nu(B)-f\mu(B)\geq 0$.

$\therefore f\in\mathcal{H}\triangleq\{f$ 非负, $\mathcal{S}$-可测, $\int_B f\,d\mu\leq\nu(B)$, $\forall B\in\mathcal{S}\}$.

$\mathcal{H}$ 有 prop: ① $h_n\uparrow h$, $h_n\in\mathcal{H}\Rightarrow h\in\mathcal{H}$.

$\because\int_B h_n\,d\mu\leq\nu(B)$ 由 MCT, $\lim\int_B h_n\,d\mu=\int_B h\,d\mu\leq\nu(B)$.

② $h_1,h_2\in\mathcal{H}\Rightarrow h_1\vee h_2=\max(h_1,h_2)\in\mathcal{H}$.

$\because\int_B h_1\vee h_2\,d\mu=\int_{B\cap\{h_1\leq h_2\}}h_2\,d\mu+\int_{B\cap\{h_2<h_1\}}h_1\,d\mu$

$\leq\nu(B\cap\{h_1\leq h_2\})+\nu(B\cap\{h_2<h_1\})=\nu(B)$.

定义 $K\triangleq\sup_{h\in\mathcal{H}}\int_S h\,d\mu\in[0,\nu(S)]$.

则 $\exists\{h_n\}\subset\mathcal{H}$, s.t. $\int_S h_n\,d\mu>K-\frac{1}{n}$

取 $f_n\triangleq\max\{h_1,\cdots,h_n\}$, 则 $f_n\uparrow$ 非负可测, $f_n\overset{②}{\in}\mathcal{H}$. 令 $f\triangleq\lim f_n\overset{①}{\in}\mathcal{H}$

由 MCT 得 $\int_S f\,d\mu=K$

第三章 条件数学期望

称 $E[X\mid\sigma(Z)]$ ($E[X\mid Z]$) 为 r.v. X 关于 $\sigma(Z)$ 的条件数学期望

~~Q: $\forall X\in L^1(\mathcal{F})$, $\forall\sigma$-代数 $\mathcal{G}\subset\mathcal{F}$, 如何得到 $E[X\mid\mathcal{G}]$?~~

Thm. (条件数学期望 $\exists 1$ 性) 设 $X\in L^1(\mathcal{F})$, σ-代数 $\mathcal{G}\subset\mathcal{F}$, 则 $\exists 1$ (a.e.意义下) r.v. $Y\in\mathcal{G}$, s.t. $E[Y\mathbb{1}_G]=E[X\mathbb{1}_G]$, $\forall G\in\mathcal{G}$. (若 $\tilde{Y}$ 也满足, 则 $\tilde{Y}\overset{a.e.}{=}Y$)

记 $E[X\mid\mathcal{G}]\triangleq Y$, 称为 r.v. X 关于 σ-代数 $\mathcal{G}$ 的条件数学期望.

Def. (相互奇异的测度 $\mu\perp\nu$) 设 $(S,\mathcal{S})$ 上 σ-有限测度 μ,ν.

若 $\exists G\in\mathcal{S}$, s.t. $\mu(G)=\nu(G^c)=0$, 则称 μ 与 ν 相互奇异, 记 $\mu\perp\nu$.

Thm. (测度的 Lebesgue 分解) 设 $(S,\mathcal{S})$ 上 σ-有限测度 μ,ν. 则对给定 μ, 有唯一 Leb 分解.

$$\nu=\nu_{ac}+\nu_s,$$

其中 $\overset{(测度)}{\nu_{ac}}\triangleq f\mu$ (f 非负, $\mathcal{S}$-可测, $f\mu(B)\triangleq\int_B f\,d\mu$)

$\nu_s\triangleq\nu-f\mu\perp\mu$.

(续 pf) 下证 $\nu_s\triangleq\nu-f\mu\perp\mu$.

定义: 正负集不能全为无穷测度

考虑 $\nu_s-\frac{1}{n}\mu$, $n\geq 1$ (符号测度 sign measure.)

Thm. (Hahn 分解) 设 ϑ 为 $(S,\mathcal{S})$ 上符号测度, 则 $\exists G\in\mathcal{S}$, s.t.

(i) $\forall A\in\mathcal{S}$, $A\subset G$, $\vartheta(A)\geq 0$ (G 为 ϑ 的正集)

(ii) $\forall B\in\mathcal{S}$, $B\subset G^c$, $\vartheta(B)\leq 0$ (G^c 为 的负集)

称 (G,G^c) 为 ϑ 的 Hahn 分解.

(本质唯一性: 若 $(\tilde{G},\tilde{G}^c)$ 也为 ϑ 的 Hahn 分解, 则 $G\triangle\tilde{G}$ 为 ϑ-null set.

由 $\nu_s-\frac{1}{n}\mu$ 为符号测度, 由 Hahn 分解: (G_n,G_n^c) 为 $\nu_s-\frac{1}{n}\mu$ 的 Hahn 分解.

则 $\forall B\in\mathcal{S}$, $\int_B(f+\frac{1}{n}\mathbb{1}_{G_n})\,d\mu=f\mu(B)+\frac{1}{n}\mu(B\cap G_n)$

$=\nu(B)-\nu_s(B)+\frac{1}{n}\mu(B\cap G_n)$

($\because\nu_s(B)-\frac{1}{n}\mu(B\cap G_n)\geq\nu_s(B\cap G_n)-\frac{1}{n}\mu(B\cap G_n)\geq 0$.)

$\leq\nu(B)$

$\Rightarrow f+\frac{1}{n}\mathbb{1}_{G_n}\in\mathcal{H}$.

$\Rightarrow\mu(G_n)=0$. 否则, $\int_S(f+\frac{1}{n}\mathbb{1}_{G_n})\,d\mu=\int_S f\,d\mu+\frac{1}{m}\mu(G_n)>\int_S f\,d\mu=K$

与 $f+\frac{1}{n}\mathbb{1}_{G_n}\in\mathcal{H}\Rightarrow\int_S(f+\frac{1}{n}\mathbb{1}_{G_n})\,d\mu\leq K$ 矛盾.

∴ $G \triangleq \bigcup_{n\ge1} G_n$，有 $\mu(G)=0$.

下证 $\nu_s(G^c)=0$　否则，若 $\nu_s(G^c)>0$，∃ n 充分大，s.t.

$$\nu_s(G^c)-\frac{1}{n}\mu(G^c)>0$$

$$5 \overset{G^c\subset G_n^c}{\Rightarrow} \nu_s(G^c)-\frac{1}{n}\mu(G^c)\le 0 \quad 矛盾.$$

∴ $\mu(G)=\nu_s(G^c)=0 \Rightarrow \nu_s \perp \mu$.

若 μ,ν 为 σ-有限测度，则 ∃ S 的划分 $\{B_n\}$, s.t. $\mu(B_n),\nu(B_n)<+\infty$.

定义 $\mu_n \triangleq \mathbb{1}_{B_n}\mu$，则 μ_n、ν_n 有限测度且 $\mu=\sum_{n\ge1}\mu_n$，$\nu=\sum_{n\ge1}\nu_n$

$\nu_n \triangleq \mathbb{1}_{B_n}\nu$ #

Rmk: Hahn 分解证明 outline:

定义 $\alpha \triangleq \sup_{G为\vartheta正集} \vartheta(G) \ge 0$（由于 $\varnothing$ 为 ϑ 的正集）

∃$\{G_n\}$ 为 ϑ 正集，s.t. $\alpha-\frac{1}{n}<\vartheta(G_n)\le\alpha$　（G_n 为 maximizing sequence）

取 $G \triangleq \bigcup_{n\ge1} G_n$，则 $G\in\mathcal{S}$，G 为 ϑ 正集，$\vartheta(G)=\alpha$

下证 G^c 为 ϑ 的负集.　否则，∃$A_0\subset G^c$, s.t. $\vartheta(A_0)>0$.

(Hahn Lemma) CLAIM: ∃ $G_0\subset A_0$, s.t. G_0 为 ϑ 正集，$\vartheta(G_0)>0$

则 $G_0\cup G$ 也为 ϑ 正集且 $\vartheta(G_0\cup G)=\alpha+\vartheta(G_0)>\alpha$ 矛盾!

Eg. 设 μ 为 $(S,\mathcal{S})$ 上有限测度，可测 $f\in L^1(d\mu)$

$\vartheta(B)\triangleq f\mu(B)\triangleq\int_B f\,d\mu$.

(i) ϑ 的 Hahn 分解？　$G=\{x: f(x)\ge0\}$.

(ii) ϑ 的 Jordan 分解？　$\vartheta^{\pm}=f^{\pm}\mu$

(iii) ϑ 的全变差？　$|\vartheta|=|f|\mu$.

Rmk. 设 ϑ 为 $(S,\mathcal{S})$ 上符号测度，由 Hahn 分解 (G,G^c)，定义 $\forall B\in\mathcal{S}$.

$\vartheta^+(B)\triangleq\vartheta(B\cap G)\ge0$

$\vartheta^-(B)\triangleq-\vartheta(B\cap G^c)\ge0$

则 $\vartheta=\vartheta^+-\vartheta^-$，$\vartheta^+\perp\vartheta^-$（$\vartheta$ 的唯一 Jordan 分解）
（由 $\vartheta^+(G^c)=\vartheta^-(G)=0$）

称 $|\vartheta|=\vartheta^++\vartheta^-$ 为符号测度的全变差

Thm (Radon-Nikodym) 设 $\nu\ll\mu$ 为 $(S,\mathcal{S})$ 上有限测度，则 ∃ 非负 $\mathcal{S}$-可测 f,

s.t. $\nu=f\mu$.

且若 $\nu=g\mu$，则 $f=g$, a.e.

Pf: 由 Lebesgue 分解. $\nu=\nu_{ac}+\nu_s$，$\nu_{ac}=f\mu$，$\nu_s\perp\mu$.

$\Rightarrow \exists G\in\mathcal{S}$, s.t. $\mu(G)=\nu_s(G^c)=0$.

由 $\nu\ll\mu$，$\nu(G)=0 \Rightarrow \nu_s(G)=0 \Rightarrow \nu_s(S)=0 \Rightarrow \nu_s\equiv0$.
（唯一性略）　$\Rightarrow \nu=\nu_{ac}=f\mu$. #

下证条件数学期望 ∃1 性:

定义 $\mu\triangleq\mathbb{P}|_{\mathcal{G}}$，即 $\forall G\in\mathcal{G}$，$\mu(G)=\mathbb{P}(G)$，则 $(\Omega,\mathcal{G},\mu)$ 为概率空间.

① 设 X 非负，定义 $\nu(G)=E[X\mathbb{1}_G]$，$\forall G\in\mathcal{G}$，则 ν 为 $\mathcal{G}$ 上有限测度 $(E[X]<\infty)$

∴ $\nu\ll\mu$.（原来的一个引理）

由 Radon-Nikodym ∃ 非负可积 $Y\in\mathcal{G}$, s.t. $\nu=Y\mu$. on $\mathcal{G}$.

即：$\forall G\in\mathcal{G}$，$E[X\mathbb{1}_G]=\nu(G)=\int_G Y\,d\mu=\int_G Y\,d\mathbb{P}=E[Y\mathbb{1}_G]$.

② 若 $X\in L^1(\mathcal{F})$，考虑 $X=X^+-X^-$

③ 唯一性：$\forall G\in\mathcal{G}$，$E[(Y-\tilde{Y})\mathbb{1}_G]=0$.

定义 $G_0\triangleq\{w:(Y-\tilde{Y})(w)>0\}\in\mathcal{G}$，则 $\mathbb{P}(G_0)=0$（原来一个引理）

$\tilde{G}_0\triangleq\{w:(\tilde{Y}-Y)(w)>0\}\in\mathcal{G}$，则 $\mathbb{P}(\tilde{G}_0)=0$

∴ $\mathbb{P}(\cancel{X}=\tilde{Y})=1-\mathbb{P}(G_0)-\mathbb{P}(\tilde{G}_0)=1$，即 $Y=\tilde{Y}$ a.e. #

Pf: (2) $\overset{\text{step 1}}{g(Y)} \in \mathcal{G}$

$\underline{\text{step 2}}$ $\forall G \in \mathcal{G}$, $E[\varphi(Y,Z)\mathbb{1}_G] = E[g(Y)\mathbb{1}_G]$

令 $X = \mathbb{1}_G$，则 $X \in \mathcal{G}$.

由 $\left.\begin{matrix} Y \in \mathcal{G} \\ Z \perp \mathcal{G} \end{matrix}\right\} \Rightarrow Z \perp (X,Y)$

$$\therefore E[\varphi(Y,Z)X] \xlongequal{\text{change of variable}} \int_{S_Y \times S_Z \times \mathbb{R}} \varphi(y,z)x\, P_{(X,Y,Z)}(dx,dy,dz)$$

$$\xlongequal{Z\perp(X,Y) \Rightarrow P_{(X,Y,Z)} = P_{(X,Y)} \times P_Z} \int \varphi(y,z)x\, P_{(X,Y)}(dx,dy)P_Z(dz)$$

$$\xlongequal{\text{Fubini}} \int_{\mathbb{R}\times S_Y} x\left(\int \varphi(y,z)P_Z(dz)\right)P_{(X,Y)}(dx,dy)$$

$$\xlongequal{\text{change of variable}} \int x\, g(y)\, P_{(X,Y)}(dx,dy)$$

$$\xlongequal{\text{change of variable}} E[Xg(Y)]$$ #

e.g. 设 $X = \sum_{i=1}^{N} \xi_i$，$\xi_i \overset{\text{i.i.d.}}{\sim}$ $E(\xi_i) = \mu \in \mathbb{R}$，$Var(\xi_i) = \sigma^2 > 0$，$N \sim Poisson(\lambda)$，$\lambda > 0$ 且与 ξ_i 相互独立.

求 $EX = ?$ $VarX = ?$

解：$EX = E\{E[X|\sigma(N)]\}$

$= E[f(N)]$，$f(N) \triangleq E[\sum_{i=1}^{N}\xi_i|\sigma(N)] \in \sigma(N)$，其中 $f(n) \triangleq E[\sum_{i=1}^{N}\xi_i | N=n]$ $_{n=0,1,2,\cdots}$ $= n\mu$

$\therefore EX = E[\mu N] = \lambda\mu$.

$EX^2 = E\{E[(\sum_{i=1}^{N}\xi_i)^2|\sigma(N)]\}$

$\because E[(\sum_{i=1}^{N}\xi_i)^2 | N=n] = E(\sum_{i=1}^{n}\xi_i)^2 = E(\sum_{i=1}^{n}\xi_i^2) + \sum_{i\neq j}E(\xi_i)E(\xi_j)$

可得 $VarX$. #

Pf: (4) $\forall G \in \mathcal{G}$. $E[X\mathbb{1}_G] = E\{E[X|\mathcal{G}]\mathbb{1}_G\}$

由 $X \geq 0$ a.e 知 LHS ≥ 0.

记 $G_0 = \{\omega \in \Omega : E[X|\mathcal{G}](\omega) < 0\}$. 下证 $\mathbb{P}(G_0) = 0$.

由 $E[X|\mathcal{G}] \in \mathcal{G} \Rightarrow G_0 \in \mathcal{G} \Rightarrow E\{E[X|\mathcal{G}]\mathbb{1}_{G_0}\} \geq 0$

由 $E[X|\mathcal{G}] < 0$ on $G_0 \Rightarrow E\{E[X|\mathcal{G}]\mathbb{1}_{G_0}\} \leq 0$

$\Rightarrow E[X\mathbb{1}_{G_0}] = 0 \overset{X\geq 0}{\Longrightarrow} \mathbb{P}(G_0) = 0$ #

§3.2 条件数学期望性质.

设 $(\Omega, \mathcal{F}, \mathbb{P})$ 下 r.v. X 以及 σ-代数 $\mathcal{G} \subset \mathcal{F}$.

(1) $X \in L^1(\mathcal{G})$，则 $E[X|\mathcal{G}] = X$. 特别，$X \equiv Const$ 时，$E[C|\mathcal{G}] = C$.

(X 可积且 $X \in \mathcal{G}$).

(2) $Y \in \mathcal{G}$，$Z \in \mathcal{F}$ 且 $Z \perp \mathcal{G}$ ($\sigma(Z)$ 与 $\mathcal{G}$ 独立)

$Y: \Omega \to S_Y$，$Z: \Omega \to S_Z$.

记 $\varphi: S_Y \times S_Z \to \mathbb{R}$ 可测 ($\varphi \in \mathcal{S}_Y \otimes \mathcal{S}_Z$) 使 $E[|\varphi(Y,Z)|] < \infty$.

则 $E[\varphi(Y,Z)|\mathcal{G}] = g(Y)$，其中 $g(y) = E[\varphi(y,Z)]$，$\forall y \in S_Y$.

特别，$\varphi(y,z) = f(y)h(z)$，f,h 可测 时，

$E[f(Y)h(Z)|\mathcal{G}] = f(Y)E[h(Z)]$

($\overset{\shortparallel}{g}(Y)$，其中 $g(y) = E[f(y)h(Z)] = f(y)E[h(Z)]$.)

特别，$f \equiv 1$ 时，$E[h(Z)|\mathcal{G}] = E[h(Z)]$

$h \equiv 1$ 时，$E[f(Y)|\mathcal{G}] = f(Y)$.

(3) $X \in L^1(\mathcal{F})$，则 $E\{E[X|\mathcal{G}]\} = EX$. (由 $E\{E[X|\mathcal{G}]\mathbb{1}_\Omega\} = E[X\mathbb{1}_\Omega]$ 得.)

(4) $X \in L^1(\mathcal{F})$，若 $X \geq 0$, a.e.，则 $E[X|\mathcal{G}] \geq 0$ a.e.

从而 $X \geq Y$, a.e. ($Y \in L^1(\mathcal{F})$)，则 $E[X|\mathcal{G}] \geq E[Y|\mathcal{G}]$.

(5) $X, Y \in L^1(\mathcal{F})$，$\alpha, \beta \in \mathbb{R}$，则 $E[\alpha X + \beta Y|\mathcal{G}] = \alpha E[X|\mathcal{G}] + \beta E[Y|\mathcal{G}]$.

Pf: $\forall G \in \mathcal{G}$，$E\{E[X|\mathcal{G}]\mathbb{1}_G\} = E[X\mathbb{1}_G]$

$E\{E[Y|\mathcal{G}]\mathbb{1}_G\} = E[Y\mathbb{1}_G]$

$\therefore E\{(\alpha E[X|\mathcal{G}] + \beta E[Y|\mathcal{G}])\mathbb{1}_G\} = \alpha E[X\mathbb{1}_G] + \beta E[Y\mathbb{1}_G] = E\{(\alpha X + \beta Y)\mathbb{1}_G\}$

$\Rightarrow E[\alpha X + \beta Y|\mathcal{G}] = \alpha E[X|\mathcal{G}] + \beta E[Y|\mathcal{G}]$ #

(6) (Tower property) $X \in L^1(\mathcal{F})$，σ-代数 $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$，则 $E\{E[X|\mathcal{G}]|\mathcal{H}\} = E[X|\mathcal{H}]$.

Pf: $Y \triangleq E[X|\mathcal{G}]$，$Z \triangleq E[Y|\mathcal{H}]$，则 $E[X\mathbb{1}_G] = E[Y\mathbb{1}_G]$，$\forall G \in \mathcal{G}$

$E[Y\mathbb{1}_H] = E[Z\mathbb{1}_H]$，$\forall H \in \mathcal{H}$

由 $\mathcal{H} \subset \mathcal{G}$，$H \in \mathcal{G} \Rightarrow E[Z\mathbb{1}_H] = E[Y\mathbb{1}_H] = E[X\mathbb{1}_H]$，$\forall H \in \mathcal{H} \Rightarrow Z = E[X|\mathcal{H}]$ #

e.g $\varphi(x)=x^2$, Jensen $\Rightarrow (EX)^2 \le E[X^2]$

Ex $X\in L^1(\mathcal{F})$，σ-代数 $\mathcal{G}\subset\mathcal{F}$，则 $X=E[Y|\mathcal{G}]+\xi$，其中 $E[\xi]=0$，$E[\xi Y]=0, \forall Y\in L^2(\mathcal{G})$

证：$\xi=X-E[X|\mathcal{G}]$

则 $E\xi=EX-E\{E[X|\mathcal{G}]\}=0$

$\forall G\in\mathcal{G}$, $E[\xi 1_G]=0$

若 $Y=1_G\in L^2(\mathcal{G})$，$E[\xi Y]=0$.

若 $Y\in L^2(\mathcal{G})$ 非负简单，$Y=\sum_{n=1}^{m} a_n 1_{G_n}$，$a_n>0$.

$E[\xi 1_G]=0 \Leftrightarrow E[\xi^+ 1_G]=E[\xi^- 1_G]$

$\Rightarrow m<\infty$ 时，$E[\xi^+ Y]=E[\xi^- Y] \Leftrightarrow E[\xi Y]=0$

$m=\infty$ 时，由MCT，$E[\xi^+ Y]=E[\xi^- Y] \Leftrightarrow E[\xi Y]=0$.

若 $Y\in L^2(\mathcal{G})$ 非负，$\exists \{Y_n\}\subset L^2(\mathcal{G})$ 非负简单，s.t. $Y_n\uparrow Y$ in Ω

由MCT，$E[\xi^+ Y]=E[\xi^- Y] \Leftrightarrow E[\xi Y]=0$.

若 $Y\in L^2(\mathcal{G})$，$Y=Y^+-Y^-$，$E[\xi Y]=E[\xi Y^+]-E[\xi Y^-]=0$.

Ex. $Y\in L^2(\mathcal{F})$，σ-代数 $\mathcal{G}\subset\mathcal{F}$. $E[|Y-\eta^*|^2]=\inf_{\eta\in L^2(\mathcal{G})} E[|Y-\eta|^2]$，$\eta^*\in L^2(\mathcal{G})$

则 $\eta^*=E[Y|\mathcal{G}]$.

证：$E|Y-\eta|^2=E|Y-E[Y|\mathcal{G}]|^2+E|E[Y|\mathcal{G}]-\eta|^2+2E[(Y-E[Y|\mathcal{G}])(E[Y|\mathcal{G}]-\eta)]$

记 $Z\triangleq E[Y|\mathcal{G}]-\eta\in L^2(\mathcal{G})$

则 $E[(Y-E[Y|\mathcal{G}])Z]=E\{E[(Y-E[Y|\mathcal{G}])Z|\mathcal{G}]\}$

$=E\{Z\cdot E[Y-E[Y|\mathcal{G}]|\mathcal{G}]\}=0$.

$\therefore E|Y-\eta|^2\ge E|Y-E[Y|\mathcal{G}]|^2$，"=" 当 $\eta^*\overset{a.e.}{=}E[Y|\mathcal{G}]$ 时成立 #.

注：条件期望的不等式及收敛定理

- Holder：$E|XY|\le (E|X|^p)^{\frac1p}\cdot(E|Y|^q)^{\frac1q}$，$1<p<+\infty$，$X\in L^p(\mathcal{F})$，$Y\in L^q(\mathcal{F})$

  特别，$E|XY|\le (E|X|^2)^{\frac12}\cdot(E|Y|^2)^{\frac12}$ (Cauchy-Schwarz)

  条件：$E[|XY||\mathcal{G}]\le (E[|X|^p|\mathcal{G}])^{\frac1p}\cdot(E[|Y|^q|\mathcal{G}])^{\frac1q}$

  (令 $H\triangleq L^2(\mathcal{F})=\{r.v. X\in\mathcal{F}: E|X|^2<\infty\}$，$\|X\|_{L^2}\triangleq(E|X|^2)^{\frac12}$

  $(H,\|\cdot\|_{L^2})$ 为Hilbert空间，内积 $\langle X,Y\rangle=E[X\cdot Y]$，$X,Y\in H$)

- Jensen：$X\in L^1(\mathcal{F})$，$\varphi:\mathbb{R}\to\mathbb{R}$ 凸 且 $\varphi(X)$ 可积，则

  $\varphi(EX)\le E[\varphi(X)]$

  条件：$\varphi(E[X|\mathcal{G}])\le E[\varphi(X)|\mathcal{G}]$.

- MCT r.v.s $\{X_n\}_{n\ge1}$ 非负(可积)↑，则 $\lim_{n\to\infty}E[X_n|\mathcal{G}]=E[\lim_{n\to\infty}X_n|\mathcal{G}]$

- Fatou r.v.s $\{X_n\}_{n\ge1}$ 非负(可积)，则 $E[\underline{\lim}X_n|\mathcal{G}]\le\underline{\lim}E[X_n|\mathcal{G}]$

- DCT. r.v.s $\{X_n\}_{n\ge1}$，$|X_n|\le Y$ a.e.，$Y\in L^1(\mathcal{F})$，则 $\lim_{n\to\infty}E[X_n|\mathcal{G}]=E[\lim_{n\to\infty}X_n|\mathcal{G}]$.

§3.3 条件数学期望的应用.

统计决策：Input $X\in\mathbb{R}^n$，Output $Y\in\mathbb{R}$.

找可测 $f:\mathbb{R}^n\to\mathbb{R}$，s.t. $f(X)$ 某种意义下逼近 Y，i.e.，$Y\approx f(X)$.

即考虑 $\inf_{f\in B(\mathbb{R}^n\to\mathbb{R})} E[L(Y,f(X))]$ $\Rightarrow f(X)\in\sigma(X)$

特别，若考虑 $\inf_{\eta\in L^2(\sigma(X))} E|Y-\eta|^2=E|Y-\eta^*|^2$，则 $\eta^*=E[Y|\sigma(X)]\in\sigma(X)$.

$\Rightarrow\exists f^*$, s.t. $\eta^*=f^*(X)$

称 $f^*(x)=E[Y|X=x]$，$x\in\mathbb{R}^n$ 为<u>回归函数</u>.

设 $(x_1^{\in\mathbb{R}^n},\cdots,x_N^{\in\mathbb{R}^n})$ 为样本 $(X_1\cdots X_N)$ 一个观测，$(y_1,\cdots,y_N)$ 为 $(Y_1,\cdots,Y_N)\cdots$

则 $\hat{y}_i=x_i^T\beta$

$RSS(\beta)=\sum_{i=1}^{N}(y_i-\hat{y}_i)^2=(y-X\beta)'(y-X\beta)$ 考虑 $\inf_{\beta\in\mathbb{R}^{n+1}} RSS(\beta)$ 得

$\beta^*=(X^TX)^{-1}X^Ty$ 最小二乘法 线性回归分析

正则条件概率分布 $E[1_A|\mathcal{G}]=P(A|\mathcal{G})$，$A\in\mathcal{F}$.

$\mu(w,A)\triangleq P(A|\mathcal{G})(w)$

(1) $\forall w\in\Omega$：$\mu(w,\cdot)$ 是否为 $\mathcal{F}$ 上概率测度？ (x)

(2) $\forall A\in\mathcal{F}$：$\mu(\cdot,A)$ 是否为 r.v.？

是否存在 $\tilde{\mu}(w,A)=\mu(w,A)$ a.s. $\forall A\in\mathcal{F}$，s.t. $\tilde{\mu}$ 满足(1)(2).？

Pf of lem: $X_n \xrightarrow{a.e.} X$. 记 $\Omega_0 \triangleq \{w\in\Omega: \lim X_n(w) = X(w)\}$, 则 $P(\Omega_0)=1$

$\forall w\in\Omega_0$, $\forall\varepsilon>0$, $\exists\, m(w)\geq 1$, s.t. $\forall n\geq m(w)$, $|X_n(w)-X(w)|<\varepsilon$.

$\therefore w\in \bigcup_{m\geq 1}\bigcap_{n\geq m}\{|X_n-X|\leq\varepsilon\} = \underline{\lim}\{|X_n-X|\leq\varepsilon\}$.

$\therefore \Omega_0\subset \underline{\lim}\{|X_n-X|\leq\varepsilon\}$

$\therefore \mathbb{P}(\underline{\lim}\{|X_n-X|\leq\varepsilon\})=1$

$\therefore \mathbb{P}(\overline{\lim}\{|X_n-X|>\varepsilon\})=0$

E.g. $X_n\xrightarrow{P}X$, $Y_n\xrightarrow{P}Y$, 则 $(X_n,Y_n)\xrightarrow{P}(X,Y)$.

证: $\forall\varepsilon>0$, $\mathbb{P}(\|(X_n,Y_n)-(X,Y)\|>\varepsilon)=\mathbb{P}(|X_n-X|^2+|Y_n-Y|^2>\varepsilon^2)$

$\leq \mathbb{P}(|X_n-X|^2>\frac{\varepsilon^2}{2}) + \mathbb{P}(|Y_n-Y|^2>\frac{\varepsilon^2}{2}) \to 0.$ #

Pf of a.e $\Rightarrow$ P: $X_n\xrightarrow{a.e.}X \Rightarrow P(\overline{\lim}\{|X_n-X|>\varepsilon\})=0$.

由Fatou: $0=P(\overline{\lim}\{|X_n-X|>\varepsilon\})\geq\overline{\lim}P(\{|X_n-X|>\varepsilon\})\geq 0$

$\Rightarrow \lim P(\{|X_n-X|>\varepsilon\})=0$ 即 $X_n\xrightarrow{P}X$. #

Pf of P $\Leftrightarrow$ a.e.子列: ($\Rightarrow$) $X_n\xrightarrow{P}X \Rightarrow \lim P(|X_n-X|>\varepsilon)=0$, $\forall\varepsilon>0$.

$\forall$子列 $\{X_{n_k}\}$, $\lim P(|X_{n_k}-X|>\varepsilon)=0$

$\therefore \forall p\geq 1$, $\exists\{X_{n_{k_p}}\}\subset\{X_{n_k}\}$, s.t. $P(|X_{n_{k_p}}-X|>\frac{1}{2^p})\leq\frac{1}{2^p}$

$\therefore \sum_{p=1}^{\infty}P(|X_{n_{k_p}}-X|>\frac{1}{2^p})<\infty \Rightarrow X_{n_{k_p}}\xrightarrow{a.e.}X$.

($\Leftarrow$) 若 $X_n\not\xrightarrow{P}X$, $\exists\varepsilon>0$, $\delta>0$, $\{X_{n_k}\}$, s.t.

$P(|X_{n_k}-X|>\varepsilon)>\delta$.

与 $\exists X_{n_{k_p}}\xrightarrow{a.e.}X \Rightarrow X_{n_{k_p}}\xrightarrow{P}X$ 矛盾. #

Ex. $Y_n\xrightarrow{P}0$, $\lim_{a\to\infty}\sup_{n\geq 1}P(|X_n|>a)=0 \Rightarrow X_nY_n\xrightarrow{P}0$.

证: $\forall\varepsilon>0$, $P(|X_nY_n|>\varepsilon)=P(|X_nY_n|>\varepsilon, |X_n|>a)+P(|X_nY_n|>\varepsilon, |X_n|\leq a)$

$\leq\sup_{n\geq 1}P(|X_n|>a)+P(|Y_n|>\frac{\varepsilon}{a})$

令 $n\to\infty$, $\overline{\lim}P(|X_nY_n|>\varepsilon)\overset{Y_n\xrightarrow{P}0}{\leq}\sup_{n\geq 1}P(|X_n|>a)+0$ 令 $a\to\infty$ 得 $\overline{\lim}P(|X_nY_n|>\varepsilon)=0$ #

第四章 r.v.列的收敛.

Def. (a.e.收敛). $X_n\xrightarrow{a.e.}X$ $(n\to\infty)$: 若 $\{X_n\}$. X 在 $(\Omega,\mathcal{F},\mathbb{P})$ 下实值 r.v.s.,

$$\mathbb{P}(\{w\in\Omega: \lim_{n\to\infty}X_n(w)=X(w)\})=1$$

Rmk. 或对映到 (S,d) 的 r.v.s. $\mathbb{P}(\{w\in\Omega: \lim_{n\to\infty}d(X_n(w),X(w))=0\})=1$.

Lem $X_n\xrightarrow{a.e.}X \Leftrightarrow \forall\varepsilon>0$, $\mathbb{P}(\{|X_n-X|>\varepsilon, \text{ i.o.}\})=0$

$\Leftrightarrow \forall\varepsilon>0$, $\mathbb{P}(\overline{\lim_{n\to\infty}}\{|X_n-X|>\varepsilon\})=0$

$\overset{B-C}{\Leftarrow} \forall\varepsilon>0$, $\sum_{n=1}^{\infty}\mathbb{P}(|X_n-X|>\varepsilon)<\infty$.

Rmk 1° (SLLN) X_i i.i.d. $EX_1=\mu\in\mathbb{R}$, 则 $\frac{1}{n}\sum X_i\xrightarrow{a.e.}\mu$, $n\to\infty$.

2° $\sum_{n=1}^{\infty}\mathbb{P}(|X_n-X|>\varepsilon)<\infty$ $(\forall\varepsilon>0)$ $\Rightarrow \lim\mathbb{P}(|X_n-X|>\varepsilon)=0$

Def. (依Prob收敛). 设 r.v.s $\{X_n\}$. X 在 $(\Omega,\mathcal{F},\mathbb{P})$ 下. 若 $\forall\varepsilon>0$,

$$\mathbb{P}(|X_n-X|>\varepsilon)\to 0 \quad (n\to\infty).$$

则称 X_n 依概率收敛到 X, 记 $X_n\xrightarrow{\mathbb{P}}X$ $(n\to\infty)$.

Rmk. $(\mathbb{R},|\cdot|)\to(S,\|\cdot\|)$ 赋范空间, $X_n\xrightarrow{\mathbb{P}}X \Leftrightarrow \forall\varepsilon>0$, $\mathbb{P}(\|X_n-X\|>\varepsilon)\to 0$.

Lem. $X_n\xrightarrow{a.e.}X \Rightarrow X_n\xrightarrow{P}X$.

Lem. $X_n\xrightarrow{P}X \Leftrightarrow$ (列紧) $\forall\{X_{n_k}\}\subset\{X_n\}$, $\exists$子列 a.e. 收敛于 X.

Def. (L^p 收敛) 设实值 r.v.s $\{X_n\}$. X in $(\Omega,\mathcal{F},\mathbb{P})$.

$X, X_n\in L^p(\mathcal{F})$, $p\geq 1$. 若 $E|X_n-X|^p\to 0$ $(n\to\infty)$,

则称 $X_n\xrightarrow{L^p}X$ $(n\to\infty)$.

Lem. (1) $X_n\xrightarrow{L^p}X \Rightarrow X_n\xrightarrow{\mathbb{P}}X$

(2) $X_n\xrightarrow{\mathbb{P}}X$ 且 $|X_n|\overset{a.e.}{\leq}Y\in L^p(\mathcal{F})$, $\Rightarrow X_n\xrightarrow{L^p}X$

Pf: (1) $\mathbb{P}(|X_n-X|>\varepsilon)\leq\frac{E|X_n-X|^p}{\varepsilon^p}\to 0$

(2) $X_n\xrightarrow{P}X \Rightarrow \forall\varepsilon>0$, $\mathbb{P}(|X_n-X|>\varepsilon)\to 0$.

记 $\tilde{\Omega}=\{|X_n|\leq Y, \forall n\}$ $\because \{|X|>Y+\varepsilon\}\cap\tilde{\Omega}\subset\{|X-X_n|>\varepsilon\}\cap\tilde{\Omega}$

$\therefore \mathbb{P}(\{|X|>Y+\varepsilon\}\cap\tilde{\Omega})\leq\mathbb{P}(|X-X_n|>\varepsilon)$ 令 $n\to\infty$, LHS $=0$

由 $\mathbb{P}(\tilde{\Omega})=1$ (check!), $\mathbb{P}(\{|X|>Y+\varepsilon\})=0 \Rightarrow |X|\leq Y$ a.e.

$\therefore E|X_n-X|^p=E|X_n-X|^pI\{|X_n-X|\leq\varepsilon\}+E|X_n-X|^pI\{|X_n-X|>\varepsilon\}$

$\leq\varepsilon^p+2^pE|Y|^pI\{|X_n-X|>\varepsilon\}$ 令 $n\to\infty$ 得 $\overline{\lim}$ LHS $\leq\varepsilon^p$ 由 ε 任意 #

Pf of 唯一性:

(1) (a.e.) $\Omega_1 \triangleq \{\lim X_n = X\}$, $\Omega_2 \triangleq \{\lim X_n = Y\}$

则 $\mathbb{P}(\Omega_1) = \mathbb{P}(\Omega_2) = 1 \Rightarrow \mathbb{P}(\Omega_1 \cap \Omega_2) = 1$.

在 $\Omega_1 \cap \Omega_2$ 上, $X = Y$. $\Rightarrow X = Y$ a.e.

(2) ($\mathbb{P}$). $\mathbb{P}(|X-Y| > \varepsilon) \le \mathbb{P}(|X - X_n| > \frac{\varepsilon}{2}) + \mathbb{P}(|X_n - Y| > \frac{\varepsilon}{2}) \to 0 \ (n\to\infty)$.

(3) (L^p). $E|X-Y|^p \le 2^{p-1}(E|X-X_n|^p + E|X_n - Y|^p) \longrightarrow 0.\ (n\to\infty)$ #.

E.g. 设 $F_n(x) \triangleq \frac{e^{nx}}{1+e^{nx}} = \frac{1}{1+e^{-nx}}$, $n\ge 1$, $x\in\mathbb{R}$, 则 F_n 是一个分布函数,

但 $\lim F_n(x) = \widetilde{F}(x) = \begin{cases} 1, & x>0 \\ \frac{1}{2}, & x=0 \\ 0, & x<0. \end{cases}$ 不是分布函数.

$\because$ 定义 $F(x) = \mathbb{1}_{[0,\infty)}(x)$ 为一分布函数, $C_F = \{x : x \ne 0\}$, $F_n \to F$, $\forall x\in C_F$

$\Rightarrow X_n \xrightarrow{d} X$

Pf of Skorohod:

定义 $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}}) = ([0,1], \mathcal{B}_{[0,1]}, m)$, m 为 Lebesgue 测度.

$\forall w\in[0,1]$, $\widetilde{X}(w) \triangleq \inf\{x\in\mathbb{R}, F(x)\ge w\}$, $F(x) = \mathbb{P}(X\le x)$

则 $w \mapsto \widetilde{X}(w)$ 单增, 是 $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ 上 r.v.

且 $\widetilde{X}(w) \le x \Longleftrightarrow F(x) \ge w$ (check!)

$\therefore \widetilde{\mathbb{P}}(\widetilde{X}\le x) = \widetilde{\mathbb{P}}(\{w\in\widetilde{\Omega} : F(x)\ge w\})$

$= m(\{w\in[0,1] : w \le F(x)\}) = F(x) = \mathbb{P}(X\le x)$

$\therefore \widetilde{X} \stackrel{d}{=} X$.

同理, $\forall w\in[0,1]$, $\widetilde{X}_n(w) \triangleq \inf\{x\in\mathbb{R}, F_n(x)\ge w\}$, 有 $\widetilde{X}_n \stackrel{d}{=} X_n$

下证 $\widetilde{X}_n \xrightarrow{a.e.} \widetilde{X}$ under $\widetilde{\mathbb{P}}$ 即可.

① $\forall w\in\widetilde{\Omega}$, $\varepsilon>0$, $\exists x\in C_F$, s.t. $\widetilde{X}(w) - \varepsilon < x < \widetilde{X}(w)$.

$\Rightarrow F(x) < w$, 又 $\because X_n \xrightarrow{d} X$ $\therefore F_n(x) \to F(x)$

$\therefore \exists M_1$, s.t. $\forall n\ge M_1$, $F_n(x) < w$. $\Rightarrow \widetilde{X}_n(w) > x > \widetilde{X}(w) - \varepsilon \Rightarrow \varliminf \widetilde{X}_n(w) \ge \widetilde{X}(w)$

(收敛唯一)

Lem. $X_n \xrightarrow[\text{② } \mathbb{P} \text{ ③ } L^p]{\text{① a.e.}} X, Y \Rightarrow X \stackrel{a.e.}{=} Y$

Def. (依分布收敛). 设实值 r.v.s $\{X_n\}$. X in $(\Omega, \mathcal{F}, \mathbb{P})$.

$F_n(x) \triangleq \mathbb{P}(X_n\le x)$, $F(x) \triangleq \mathbb{P}(X\le x)$, $\forall x\in\mathbb{R}$.

记 $X_n \xrightarrow{d} X$, 若 $F_n(x) \xrightarrow{n\to\infty} F(x)$, $\forall x\in C_F = \{F$ 的连续点全体$\}$.

Rmk. 若 $\{X_n\}$. X 取值为 S (拓扑空间), 则设 $\mathcal{P}_{X_n}$, $\mathcal{P}_X$ 为其分布(概率测度)

则依分布收敛条件为 $\mathcal{P}_{X_n} \Rightarrow \mathcal{P}_X$ (概率测度弱收敛).

Thm. $X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$

$X_n \xrightarrow{P} C \Leftrightarrow X_n \xrightarrow{d} C$.

§4.2 Skorohod 表示定理

Thm. (Skorohod) r.v.s $\{X_n\}$. X 在 $(\Omega, \mathcal{F}, \mathbb{P})$ 下实值, $X_n \xrightarrow{d} X$, 则

$\exists (\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ 下一列 r.v. $\{\widetilde{X}_n\}$. $\widetilde{X}$, s.t.

$\widetilde{X}_n \stackrel{d}{=} X_n$, $\forall n\ge 1$

$\widetilde{X} \stackrel{d}{=} X$

$\widetilde{X}_n \xrightarrow{a.e.} \widetilde{X}$ under $\widetilde{\mathbb{P}}$ ($\widetilde{\mathbb{P}}(\lim_{n\to\infty} \widetilde{X}_n = \widetilde{X}) = 1$.)

Rmk. 若 $\{X_n\}$, X 取值为 (S, d) (此空间为 Polish 空间, 完备可分)

Skorohod 表示定理仍成立. (Durrett: Stochastic Calculus.)

② $\forall w' > w$, $\varepsilon>0$, $\exists y\in C_F$, s.t. $\widetilde{X}(w') < y < \widetilde{X}(w') + \varepsilon$.

$\Rightarrow F(y) \ge w' > w$ 又 $\because X_n \xrightarrow{d} X$ $\therefore F_n(y) \to F(y)$

$\therefore \exists M_2$, s.t. $\forall n\ge M_2$, $F_n(y) > w \Rightarrow \widetilde{X}_n(w) \le y < \widetilde{X}(w') + \varepsilon$.

$\Rightarrow \varlimsup \widetilde{X}_n(w) \le \widetilde{X}(w')$

综上 $\widetilde{X}(w) \le \varliminf \widetilde{X}_n(w) \le \varlimsup \widetilde{X}_n(w) \le \widetilde{X}(w')$

$\forall w\in C_{\widetilde{X}}$. 再令 $w' \downarrow w$, 则 $\lim \widetilde{X}_n(w) = \widetilde{X}(w)$.

$\widetilde{\mathbb{P}}(C_{\widetilde{X}}^c) = m(C_{\widetilde{X}}^c) = 0 \Rightarrow \widetilde{\mathbb{P}}(\lim \widetilde{X}_n = \widetilde{X}) = 1 \Rightarrow \widetilde{X}_n \xrightarrow{a.e.} \widetilde{X}$ under $\widetilde{\mathbb{P}}$. #

Pf of 连续映射:

(i) $X_n \xrightarrow{a.e.} X \Rightarrow g(X_n) \xrightarrow{a.e.} g(X)$

$\because \{X \notin D_g\} \cap \{\lim X_n = X\} \subset \{\lim g(X_n) = g(X)\}$

$\because \forall w \in \{X \notin D_g\} \cap \{\lim X_n = X\}$, $\lim g(X_n) = g(X)$

$\therefore \mathbb{P}(\lim g(X_n) = g(X_n)) \geq \mathbb{P}(X \notin D_g, \lim X_n = X)$

$= \mathbb{P}(\lim X_n = X) - \mathbb{P}(\lim X_n = X \in D_g)$

$\therefore g(X_n) \xrightarrow{a.e.} g(X)$ $\quad = \mathbb{P}(\lim X_n = X) = 1$

(ii) 定义 $B_\delta^\varepsilon \triangleq \{X \notin D_g : \exists y, s.t. |x-y| < \delta$ 但 $|g(x)-g(y)| > \varepsilon\}$

则 $B_\delta^\varepsilon \downarrow \phi$, as $\delta \downarrow 0$

$\therefore \mathbb{P}(|g(X_n) - g(X)| > \varepsilon) \leq \underbrace{\mathbb{P}(X \in B_\delta^\varepsilon)}_{\to 0\,(\delta \to 0)} + \underbrace{\mathbb{P}(X \in D_g)}_{=0} + \underbrace{\mathbb{P}(|X_n - X| > \delta)}_{\to 0\,(n \to \infty)}$

$\therefore g(X_n) \xrightarrow{\mathbb{P}} g(X)$

(iii) $X_n \xrightarrow{d} X$, 由 Skorokhod 表示 thm, $\exists (\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$, $\{\tilde{X}_n\}$, $\tilde{X}$, s.t.

$\tilde{X}_n \overset{d}{=} X_n$, $\tilde{X} \overset{d}{=} X$, $\tilde{X}_n \xrightarrow{a.e.} \tilde{X}$ under $\tilde{\mathbb{P}}$.

由(i) 及 $\tilde{\mathbb{P}}(\tilde{X} \in D_g) \overset{\tilde{X} \overset{d}{=} X}{=} \mathbb{P}(X \in D_g) = 0$ 知 $g(\tilde{X}_n) \xrightarrow{a.e.} g(\tilde{X})$ under $\tilde{\mathbb{P}}$.

即, $\forall x \in C_{\tilde{F}_g}$, 记 $\tilde{F}_g(x) \triangleq \tilde{\mathbb{P}}(g(\tilde{X}) \leq x)$, $\tilde{F}_{g,n}(x) \triangleq \tilde{\mathbb{P}}(g(\tilde{X}_n) \leq x)$, 有

$\tilde{F}_{g,n}(x) \to \tilde{F}_g(x)$ $(n \to \infty)$

另外, $\tilde{F}_g(x) = \mathbb{P}(g(X) \leq x) \triangleq F_g(x)$ ($\because g(\tilde{X}) \overset{d}{=} g(X)$.)

$\tilde{F}_{g,n}(x) = \mathbb{P}(g(X_n) \leq x) \triangleq F_{g,n}(x)$

$\Rightarrow C_{\tilde{F}_g} = C_{F_g} \Rightarrow F_{g,n}(x) \to F_g(x)$ $(n \to \infty)$, $\forall x \in C_{F_g}$.

$\therefore g(X_n) \xrightarrow{d} g(X)$ #

Thm. (连续映射定理) 设 $g: \mathbb{R} \to \mathbb{R}$ 可测, $D_g = \{x \in \mathbb{R}, x$ 为 g 的不连续点$\}$

若 $\mathbb{P}(X \in D_g) = 0$, 则.

$X_n \xrightarrow[\mathbb{P}/d]{a.e.} X \Rightarrow g(X_n) \xrightarrow[\mathbb{P}/d]{a.e.} g(X)$

Rmk. 1° $\left.\begin{array}{l} X_n \xrightarrow{\mathbb{P}} X \\ Y_n \xrightarrow{\mathbb{P}} Y \end{array}\right\} \overset{CMT}{\Longrightarrow} \left\{\begin{array}{l} X_n + Y_n \xrightarrow{\mathbb{P}} X + Y \\ X_n \cdot Y_n \xrightarrow{\mathbb{P}} X \cdot Y \\ X_n / Y_n \xrightarrow{\mathbb{P}} X / Y \end{array}\right.$

$\because (X_n, Y_n) \xrightarrow{\mathbb{P}} (X, Y)$, 取 $g(x,y) = x+y$ 或 xy 或 x/y, 由CMT得.

2° $\left.\begin{array}{l} X_n \xrightarrow{d} X \\ Y_n \xrightarrow{d} C \end{array}\right. \Rightarrow (X_n, Y_n) \xrightarrow{d} (X, C)$, C 为常数 (slutsky)

§4.3 概率测度的弱收敛 (weak convergence)

Def. 设 S: 拓扑空间, $\mathcal{B}_S$ 为 S 上 Borel-σ代数, 概率测度 $\{\mu_n\}$ 和 μ 在 $(S, \mathcal{B}_S)$ 上

($\mu_n \Rightarrow \mu$) 若 $\mu_n(f) \to \mu(f)$, $\forall f \in C_b(S)$.

其中 $\mu_n(f) = \int f d\mu_n = \langle \mu_n, f \rangle$, $\mu(f) = \int f d\mu = \langle \mu, f \rangle$. 则称 μ_n 弱收敛到 μ. 记为 $\mu_n \Rightarrow \mu$.

(Riesz 表示 thm: $\forall C[0,1] = C_b[0,1]$ 上连续线性泛函 $\varphi(f) \overset{\exists \mu}{=} \int f d\mu$,

其中 $\mu \in M[0,1] = \{[0,1]$ 上所有有限的符号Borel测度, 有有限全变差$\}$)

Def. (Borel 概率测度) 设 S: 拓扑空间, $\mathcal{B}_S$ 为 Borel-σ-代数.

称 μ 为 Borel 概率测度, 若 μ 为 $\mathcal{B}_S$ 上概率测度

记 $\mathcal{P}(S) = \{\mu: \mathcal{B}_S \mapsto [0,1] : \mu$ 为概率测度$\}$

Rmk. 1° 实数列 $\{x_n\} \to x \in \mathbb{R}$, $\delta_{x_n}, \delta_x \in \mathcal{P}(\mathbb{R}) \Rightarrow \delta_{x_n} \Rightarrow \delta_x$, $n \to \infty$.

Pf: $\delta_{x_n}(f) = \int_{\mathbb{R}} f(y)\, \delta_{x_n}(dy) = f(x_n)$, $\forall f \in C_b(\mathbb{R})$

$\delta_x(f) = f(x)$.

则 $\delta_{x_n}(f) \to \delta_x(f) \Rightarrow \delta_{x_n} \Rightarrow \delta_x$

2° S-值 r.v.s $\{X_n\}$, X in $(\Omega, \mathcal{F}, \mathbb{P})$, 则 $\mathcal{P}_{X_n}, \mathcal{P}_X \in \mathcal{P}(S)$ ($\mathcal{P}_{X_n} = (X_n)_\# \mathbb{P}$, $\mathcal{P}_{X_n}(B) = \mathbb{P}(X_n^{-1}(B))$)

$\mathcal{P}_{X_n} \Rightarrow \mathcal{P}_X \Leftrightarrow Ef(X_n) \to Ef(X)$, $\forall f \in C_b(S) \Leftrightarrow X_n \Rightarrow X \overset{\triangle}{\Leftrightarrow} X_n \xrightarrow{d} X$

Hint for pf of lemma:

$\forall x\in S$，定义 $0\le f_k(x)\triangleq\inf\limits_{y\in S}\{f(y)+kd(x,y)\}\le f(x)$, $k\ge 1$

则 $|f_k(x)-f_k(y)|\le Ld(x,y)$, $\forall x,y\in S$.

Pf of "$\xrightarrow{d}$" $\Leftrightarrow$ "$\Rightarrow$":

("$\Rightarrow$") 若 $X_n\xrightarrow{d}X$，由 Skorohod 表示定理，$\exists(\tilde\Omega,\tilde{\mathcal F},\tilde{\mathbb P})$, $\tilde X_n,\tilde X$ s.t.

$\tilde X_n\overset{d}{=}X_n$, $\tilde X\overset{d}{=}X$, $\tilde X_n\xrightarrow{a.e.}\tilde X$ under $\tilde{\mathbb P}$

$\forall f\in C_b(\mathbb R)$，由 BCT

$$\tilde E[f(\tilde X_n)]\to\tilde E[f(\tilde X)].$$
$$\overset{\|}{E[f(X_n)]}\to\overset{\|}{E[f(X)]}\qquad\therefore P_{X_n}\Rightarrow P_X.$$

("$\Leftarrow$") 若 $P_{X_n}\Rightarrow P_X$，则 $\forall f\in C_b(\mathbb R)$, $E[f(X_n)]\to E[f(X)]$.

$\because F_n(x)=E[1_{(-\infty,x]}(X_n)]$

$F(x)=E[1_{(-\infty,x]}(X)]$.

$\forall\alpha\in\mathbb R$，$\exists$ 非负有界 lip $\{f_k^{\pm}\}$, s.t. $f_k^+(x)\downarrow 1_{(-\infty,\alpha]}(x)$

$f_k^-(x)\uparrow 1_{(-\infty,\alpha)}(x)$.

两边取期望得 $E[f_k^-(X_n)]\le\mathbb P(X_n<\alpha)$

令 $n\to\infty$ 得 $E[f_k^-(X)]\le\varliminf\limits_{n\to\infty}\mathbb P(X_n<\alpha)$

令 $k\to\infty$ 得 (MCT) $E[1_{(-\infty,\alpha)}(X)]\le\varliminf\limits_{n\to\infty}\mathbb P(X_n<\alpha)$

$F(\alpha-)\le\varliminf F_n(\alpha-)$

同理，$f_k^+(X_n)\ge 1_{(-\infty,\alpha]}(X_n)$

取期望，令 $n\to\infty$，$E[f_k^+(X)]\ge\varlimsup\limits_{n\to\infty}F_n(\alpha)$

令 $k\to\infty$，$F(\alpha)\ge\varlimsup F_n(\alpha)$

$\forall\alpha\in C_F$，$F_n(\alpha-)=F_n(\alpha)\Rightarrow\lim F_n(\alpha)=F(\alpha)\Rightarrow X_n\xrightarrow{d}X$ #

Lem. 度量空间 (S,d)，函数 $f:S\to\mathbb R_{\ge0}$ l.s.c. ($f(x)\le\varliminf\limits_{y\to x}f(y)$ 或 $\{f(x)\le\alpha\}$ 闭.)

则 $\exists$ 非负 Lip 函数（一致连续）f_k,

s.t. $f_k(x)\uparrow f(x)$, $\forall x\in\mathbb R$.

Rmk：1_C, $C\overset{开}{\subset}S$ 是 l.s.c. 非负有界，则 $\exists$ 非负有界 Lip 函数 $\{f_k\}\uparrow 1_C$ 函数

Lem. $D\subset S$ 闭，$\exists$ 非负有界 Lip 函数 $\{g_k\}$, s.t. $g_k(x)\downarrow 1_D(x)$, $\forall x\in S$.

Pf：$\forall x\in S$，$x\mapsto d(x,D)=\inf\limits_{y\in S}d(x,y)$ 是 Lip 函数

由 D 闭，$d(x,D)=0\Leftrightarrow x\in D$.

$g_k(x)\triangleq\{1-\frac{d(x,D)}{\frac1k}\}^+=\{1-kd(x,D)\}^+$ 满足. #.

Thm. $\{X_n\}$, X 为 $(\Omega,\mathcal F,\mathbb P)$ 下实值 r.v.s，则 $X_n\xrightarrow{d}X\Leftrightarrow P_{X_n}\Rightarrow P_X$.

$(F_n(x)\to F(x),\forall x\in C_F)$ $(\forall f\in C_b(\mathbb R), E[f(X_n)]\to E[f(X)])$

Rmk. Character function

X_n C.F. $\Phi_{X_n}(\theta)\triangleq E[e^{i\theta X_n}]$, $\theta\in\mathbb R$,

X C.F. $\Phi_X(\theta)\triangleq E[e^{i\theta X}]$, $\theta\in\mathbb R$.

$X_n\xrightarrow{d}X\Rightarrow\Phi_{X_n}(\theta)\to\Phi_X(\theta)$, $\forall\theta\in\mathbb R$

Pf：$X_n\xrightarrow{d}X\Leftrightarrow E[f(X_n)]\to E[f(X)]$, $\forall f\in C_b(\mathbb R)$

取 $f(x)=\cos(\theta x)$, $\sin\theta(x)$，则

$E[\cos(\theta X_n)]\to E[\cos(\theta X)]\Rightarrow Ee^{i\theta X_n}\to Ee^{i\theta X}$.

$E[\sin(\theta X_n)]\to E[\sin(\theta X)]$ #.

Thm (Portmanteau) 度量空间 (S,d)，$\mu_n,\mu\in\mathcal P(S)$，以下等价：

(1) $\mu_n\Rightarrow\mu$ $(\forall f\in C_b(S),\ \mu_n(f)\to\mu(f))$

(2) $\forall f\in UC_b(S)\triangleq\{f:S\to\mathbb R$ 一致连续有界$\}$, $\mu_n(f)\to\mu(f)$

(3) $D\subset S$ 闭，$\varlimsup\mu_n(D)\le\mu(D)$

(4) $C\subset S$ 开，$\mu(C)\le\varliminf\mu_n(C)$

(5) $\forall A\in\mathcal B_S$ 是 μ-连续集，$\lim\limits_{n\to\infty}\mu_n(A)=\mu(A)$

$\Leftrightarrow\mu(\partial A)=0$, $\partial A=\bar A\setminus A^\circ$

Rmk (2') $\forall f\in Lip_b(S)\triangleq\{f:S\to\mathbb R$ 有界 lip$\}$, $\mu_n(f)\to\mu(f)$.

Pf of portmanteau: (1) ⇒ (2) ⇒ (2') ✓ (3) ⇔ (4) ✓

$\because$ $Lip_b(S) \subset UC_b(S) \subset C_b(S)$.

(2') ⇒ (3): 由 Lem, $\exists \{f_k\} \subset Lip_b(S)$, $\downarrow 1_D$ 逐点, $\mu(D) = \mu(1_D) \overset{MCT}{=} \lim\limits_{k\to\infty} \int f_k d\mu$

$= \lim\limits_{k\to\infty} (\lim\limits_{n\to\infty} \int f_k d\mu_n)$

$\geq \cancel{\lim\limits_{k\to\infty}} (\overline{\lim\limits_{n\to\infty}} \int 1_D d\mu_n)$

$= \overline{\lim} \mu_n(D)$.

(3)+(4) ⇒ (5): $\forall A$ 为 μ-连续集, $\mu(\bar{A}) = \mu(A^\circ)$

$\mu(A^\circ) \overset{(4)}{\leq} \underline{\lim} \mu_n(A^\circ) \leq \overline{\lim} \mu_n(\bar{A}) \overset{(3)}{\leq} \mu(\bar{A})$

$\therefore \lim \mu_n(A) = \mu(A)$.

(5) ⇒ (1) ① $S = \mathbb{R}$ 时, $A = (-\infty, \alpha]$ 为 μ-连续集

$\forall \alpha \in C_F$, $F(\alpha) = F(\alpha-)$ ⇒ $\mu(\partial A) = 0$, A 为 μ-连续集

$\overset{(5)}{\Rightarrow} \forall \alpha \in C_F$, $\lim\limits_{n\to\infty} F_n(\alpha) = F(\alpha)$

$\Rightarrow F_n \Rightarrow F \iff \mu_n \Rightarrow \mu$

$(X_n \xrightarrow{d} X)$ $(P_{X_n} \Rightarrow P_X)$

② (S, d) 一般时, $\forall f \in C_b(S)$, $\|f\|_\infty < \infty$.

定义 push forward prob. measure

$\nu \triangleq f_\# \mu$ 为概率测度.

即 $\forall B \in \mathcal{B}_{\mathbb{R}}$, $\nu(B) = \mu(\{f \in B\})$.

且 $\forall a < -\|f\|_\infty \leq \|f\|_\infty < b$, $\nu((a,b)^c) = 0$

$\forall \varepsilon > 0$, $\exists$ $a = S_0 < \cdots < S_m = b$ $(m \geq 2)$, s.t.

(i) $S_i - S_{i-1} < \varepsilon$

(ii) $\nu(\{S_i\}) = 0$, 由 ν 有限测度

取 $A_i := f^{-1}[S_{i-1}, S_i)$, $i = 1, \cdots, m$, 则 $A_i \cap A_j = \phi$ $(\forall i \neq j)$

$S = \bigcup\limits_{i=1}^{m} A_i$, $\mu(\partial A_i) = \nu(\{S_{i-1}\}) + \nu(\{S_i\}) = 0$

$\therefore A_i$ 为 μ-连续集

$\overset{由(5)}{\Longrightarrow}$ $\mu_n(A_i) \to \mu(A_i)$, $i = 1, \cdots, m$

定义 $h(x) = \sum\limits_{i=1}^{m} S_{i-1} 1_{A_i}(x)$, $\forall x \in S$ 为简单函数

$h(x) \leq f(x) = \sum\limits_{i=1}^{m} f(x) 1_{[S_{i-1}, S_i)}(f(x)) < \sum S_i 1_{A_i}(x) = \sum (S_i - S_{i-1} + S_{i-1}) 1_{A_i}(x)$

$= h(x) + \varepsilon$

$\therefore \forall x \in S$, $h(x) \leq f(x) < h(x) + \varepsilon$.

$\therefore |\mu_n(f) - \mu(f)| \leq |\mu_n(f-h)| + |\mu_n(h) - \mu(h)| + |\mu(f-h)|$

$\leq \mu_n(|f-h|) + |\mu_n(h) - \mu(h)| + \mu(|f-h|)$

$\leq \varepsilon \cdot \mu_n(S) + \varepsilon \cdot \mu(S) + \underbrace{\sum\limits_{i=1}^{m} |S_{i-1}| |\mu_n(A_i) - \mu(A_i)|}_{\to 0 (n\to\infty)}$

$\therefore \overline{\lim} |\mu_n(f) - \mu(f)| \leq 2\varepsilon$

由 ε 任意性, $\mu_n(f) \to \mu(f)$. #

注: $X_n \xrightarrow{d} X \iff \forall f \in C_b(S)$, $P_{X_n}(f) \to P_X(f) \iff \forall f \in C_b(S)$, $E[f(X_n)] = E[f(X)]$

<u>Thm.</u> (Portmanteau, r.v. 形式) 度量空间 (S, d), $\{X_n\}_{n\geq 1}$, X 为 $(\Omega, \mathcal{F}, \mathbb{P})$ 下 r.v.s, 则以下等价.

(1) $X_n \xrightarrow{d} X$, 即 $\forall f \in C_b(S)$, $Ef(X_n) \to Ef(X)$

(2) $\forall f \in UC_b(S)$ 或 $\forall f \in Lip_b(S)$, $Ef(X_n) \to Ef(X)$

(3) $\forall D \subset S$ 闭, $\overline{\lim} \mathbb{P}(X_n \in D) \leq \mathbb{P}(X \in D)$

(4) $\forall C \subset S$ 开, $\mathbb{P}(X \in C) \leq \underline{\lim} \mathbb{P}(X_n \in C)$

(5) $\forall A \in \mathcal{B}_S$, $\mathbb{P}(X \in \partial A) = 0$, 有 $\lim\limits_{n\to\infty} \mathbb{P}(X_n \in A) = \mathbb{P}(X \in A)$

Ex. 实值 $\left.\begin{array}{l} |X_n - Y_n| \xrightarrow{\mathbb{P}} 0 \\ X_n \xrightarrow{d} X \end{array}\right\} \Rightarrow Y_n \xrightarrow{d} X$.

证: 由 Portmanteau, 只须证 $\forall f \in Lip_b(\mathbb{R})$, $Ef(Y_n) \to Ef(X)$

$|Ef(Y_n) - Ef(X)| \leq |Ef(Y_n) - Ef(X_n)| + \underbrace{|Ef(X_n) - Ef(X)|}_{\xrightarrow{X_n \xrightarrow{d} X,\ portmanteau} 0\ (n\to\infty)}$

又由 $|Ef(Y_n) - Ef(X_n)|$

$\leq E|f(Y_n) - f(X_n)|$

$\leq \|f\|_{Lip} E|Y_n - X_n| 1_{|Y_n - X_n| \leq \varepsilon} + \underbrace{2\|f\|_\infty P(|Y_n - X_n| > \varepsilon)}_{\to 0 (n\to\infty)}$

$\leq \|f\|_{Lip} \varepsilon$

得 $|Ef(Y_n) - Ef(X)| \xrightarrow{n\to\infty} 0$ #

Pf of Slutzky: 只须证 $\forall f\in C_b(\mathbb{R}^2)$, $Ef(X_n,Y_n)\to Ef(X,C)$

Step 1 $(X_n,C)\xrightarrow{d}(X,C)$ $\Leftrightarrow$ $Ef(X_n,C)\to Ef(X,C)$

∵ 定义 $g(x)\triangleq f(x,C)$, 则 $g\in C_b(\mathbb{R})$

由 $X_n\xrightarrow{d}X$, $Eg(X_n)\to Eg(X)$

Step 2. $|(X_n,Y_n)-(X_n,C)|\xrightarrow{\mathbb{P}}0$.

∵ $=|Y_n-C|\xrightarrow[\mathbb{P}]{∵Y_n\xrightarrow{\mathbb{P}}C}0$

由 step 1.2 及上页右例 Ex. $(X_n,Y_n)\xrightarrow{d}(X,C)$. #.

Pf of "Prob. measure in Polish space is tight."

由 S 可分, ∃可数稠密 $D=\{d_1,d_2\}$, 则 $S=\bigcup_{i=1}^{\infty}B_\delta(d_i)$, $\forall\delta>0$.

$\forall\varepsilon>0,\forall m\geq1$, $\exists n_m\geq1$, s.t. $\mu\left(\bigcup_{i=1}^{n_m}B_{\frac{1}{m}}(d_i)\right)>1-\frac{\varepsilon}{2^m}$ 定住

取 $K\triangleq\bigcap_{m=1}^{\infty}\bigcup_{i=1}^{n_m}\overline{B_{\frac{1}{m}}}(d_i)$ 闭集, 且 $K\subset\bigcup_{i=1}^{n_m}B_\delta(d_i)$ $(\delta>\frac{1}{m})$ $\Rightarrow$ K有界

由 S 完备, K 紧.

且 $\mu(K^c)=\mu\left(\bigcup_{m=1}^{\infty}\left(\bigcup_{i=1}^{n_m}\overline{B_{\frac{1}{m}}}(d_i)\right)^c\right)$

$\leq\sum_{m=1}^{\infty}\mu\left(\left(\bigcup_{i=1}^{n_m}B_{\frac{1}{m}}(d_i)\right)^c\right)$

$\leq\sum_{m=1}^{\infty}\frac{\varepsilon}{2^m}=\varepsilon$. #.

E.g. Polish空间 (S,d), $\mu_n\in\mathcal{P}(S)$, $n=1,\cdots,N$. 则 $\{\mu_n\}_{n=1}^{N}$ 一致胎紧.

∵ $\exists K^{(n)}\overset{紧}{\subset}S$, s.t. $\mu_n((K^{(n)})^c)\leq\varepsilon$.

取 $K=\bigcup_{n=1}^{N}K^{(n)}\overset{紧}{\subset}S$, 则 $\sup_{1\leq n\leq N}\mu_n(K^c)=\sup_{1\leq n\leq N}\mu_n((K^{(n)})^c)\leq\varepsilon$.

E.g. $\{\mu_n\}\subset\mathcal{P}(\mathbb{R})$. 若 $\lim_{M\to\infty}\sup_{n\geq1}\mu_n([-M,M]^c)=0$, 则 $\{\mu_n\}$ 一致胎紧

∵ $\exists M_0$, s.t. $\sup_{n\geq1}\mu_n([-M_0,M_0]^c)\leq\varepsilon$

E.g. $\{\mu_n\}\subset\mathcal{P}(\mathbb{R}^n)$. 若 $\lim_{M\to\infty}\sup_{n\geq1}\mu_n(\underbrace{B_M(0)^c}_{=\{|x|>M\}})=0$, 则 $\{\mu_n\}$ 一致胎紧

Thm (Slutzky) 设 $(\Omega,\mathcal{F},\mathbb{P})$ 下实值 r.v.s $\{X_n\}$, $\{Y_n\}$,

$$\left.\begin{array}{l}X_n\xrightarrow{d}X\\ Y_n\xrightarrow{d}C\end{array}\right\}\Rightarrow(X_n,Y_n)\xrightarrow{d}(X,C).$$

进一步, 由连续映射定理, $X_n+Y_n\xrightarrow{d}X+C$

$X_n\cdot Y_n\xrightarrow{d}X\cdot C$

$X_n/Y_n\xrightarrow{d}X/C$ $(C\neq0)$.

§4.4 概率测度的胎紧性 (tightness)

Def. (tightness) 设拓扑空间 S, $\mu\in\mathcal{P}(S)$. 若 $\forall\varepsilon>0$, $\exists K=K_\varepsilon\overset{紧}{\subset}S$, s.t.

$\mu(K^c)\leq\varepsilon$ (或 $\mu(K)>1-\varepsilon$)

则称 μ 为 tight (胎紧)

Rmk. 1° $\forall\mu\in\mathcal{P}(\mathbb{R})$ 是 tight 的. ∵ 概率测度连续性 $\lim_{M\to\infty}\mu([-M,M]^c)=0$.

2° $\forall\mu\in\mathcal{P}(\mathbb{R}^n)$ 是 tight 的. ∵ $\lim_{M\to\infty}\mu(\{|x|>M\})=0$.

Thm. 若 (S,d) 是完备可分度量空间 (Polish 空间), 则 $\forall\mu\in\mathcal{P}(S)$ tight.

Def. (uniform tightness) 设拓扑空间 S, $\mathcal{M}\subset\mathcal{P}(S)$. 若 $\forall\varepsilon>0$, $\exists K=K_\varepsilon\overset{紧}{\subset}S$, s.t.

$\sup_{\mu\in\mathcal{M}}\mu(K^c)\leq\varepsilon$

则称 $\mathcal{M}$ 为一致胎紧.

Rmk. 1° $\{\mu_n\}\subset\mathcal{P}(S)$ 一致胎紧, 若 $\sup_{n\geq1}\mu_n(K^c)\leq\varepsilon$.

2° 设 $\{X_n\}$ 为 $(\Omega,\mathcal{F},\mathbb{P})$ 下 S-值 r.v.s.

若 $\{\mathcal{P}_{X_n}\}\subset\mathcal{P}(S)$ 一致胎紧, 则称 $\{X_n\}$ 一致胎紧.

△ Ex. (一致 L^1-有界 ⇒ 一致胎紧). 若 $\sup_{n\geq1}E|X_n|<\infty$, 则 $\{X_n\}$ 一致胎紧, X_n: $\mathbb{R}^{n_0}$ 值

∵ 须证 $\{\mathcal{P}_{X_n}\}\subset\mathcal{P}(\mathbb{R}^{n_0})$ 一致胎紧.

记 $K_M=\{|x|\leq M\}$. 则 $\sup_{n\geq1}\mathcal{P}_{X_n}(K_M^c)=\sup_{n\geq1}\int_{\mathbb{R}^{n_0}}\mathbb{1}_{K_M^c}(x)\,\mathcal{P}_{X_n}(x)$

$\overset{换元}{=}\sup_{n\geq1}\int_\Omega\mathbb{1}_{K_M^c}(X_n(\omega))\,\mathbb{P}(X_n)$

$=\sup_{n\geq1}\mathbb{P}(|X_n|>M)\leq\sup_{n\geq1}\frac{1}{M}E|X_n|$

$=\frac{1}{M}\sup_{n\geq1}E|X_n|\to0$.

E.g. 记 $C([0,1]:\mathbb{R}) \triangleq \{$连续 $f:[0,1]\to\mathbb{R}\}$，则 $C([0,1]:\mathbb{R}) = C_b([0,1]:\mathbb{R})$

$\|f\|_\infty \triangleq \sup_{t\in[0,1]} |f(t)|$.（一致范数/最大值范数.）

$d_\infty(f,g) \triangleq \|f-g\|_\infty$.

则 $(C([0,1]:\mathbb{R}), d_\infty)$ 是 polish 空间

定义 $|f|_{0,\alpha} \triangleq \|f\|_\infty + \sup_{t\neq s} \frac{|f(t)-f(s)|}{|t-s|^\alpha}$，$\alpha\in(0,1]$.

$C^{0,\alpha}([0,1]:\mathbb{R}) \triangleq \{f\in C([0,1]:\mathbb{R}) : |f|_{0,\alpha}<\infty\}$ 称为 α-Hölder 连续函数空间

请证明：若 $\{X_n\}$ 为 $(\Omega,\mathcal{F},\mathbb{P})$ 上一列 $C([0,1]:\mathbb{R})$-值 r.v.s 且

$$\lim_{M\to\infty}\sup_{n\geq1} \mathbb{P}(|X_n|_{0,\alpha}>M) = 0,\ \exists\alpha\in(0,1]$$

则 $\{X_n\}_{n\geq1}$ 一致胎紧.

Pf: $K_M \triangleq \{f\in C([0,1]:\mathbb{R}) : |f|_{0,\alpha}\leq M\}$.，由 Arzela-Ascoli，$\overline{K_M} \overset{紧}{\subset} C([0,1]:\mathbb{R})$

$\because$ (i) $\|f\|_\infty\leq M \Rightarrow K_M(t)$ 有界

(ii) $|f(t)-f(s)|\leq M\cdot|t-s|^\alpha \Rightarrow$ 等度连续.

$\therefore \sup_{n\geq1} P_{X_n}(\overline{K_M}^c) \leq \sup_{n\geq1} P_{X_n}(K_M^c) \overset{换元}{=} \sup_{n\geq1}\mathbb{P}(|X_n|_{0,\alpha}>M) \to 0$

$\therefore \{X_n\}$ 一致胎紧 #.

E.g. $\forall p\geq1$，$\{X_n\}$ 为 $(\Omega,\mathcal{F},\mathbb{P})$ 上一列 $L^p([0,1]:\mathbb{R})$-值 r.v.s 且

(i) $\sup_{n\geq1} E\int_0^1|X_n| < \infty$

(ii) $\exists C,\alpha>0,\ \delta\in(0,1)$，s.t.

$$\sup_{n\geq1} E\left[\sup_{h\in(0,\delta)}\int_0^{1-h}|X_n(t+h)-X_n(t)|^p dt\right]\leq C\delta^\alpha$$

则 $\{X_n\}$ 一致胎紧.

Pf: $\forall \tilde{c},\tilde{\alpha},\tilde{\delta}\in(0,1)$，$\tilde{M}>0$，令 $\tilde{K}\triangleq K_{\tilde{c},\tilde{\alpha},\tilde{\delta},\tilde{M}} \triangleq \{f\in L^p([0,1]:\mathbb{R}) : \int_0^1|f|\leq\tilde{M},\ \sup_{h\in(0,\tilde{\delta})}\int_0^{1-h}|f(t+h)-f(t)|^p dt\leq\tilde{c}\tilde{\delta}^{\tilde{\alpha}}\}$

则由右侧 Thm，$\tilde{K} \subset L^p([0,1]:\mathbb{R})$ 相对紧

$\because$ (i) $\forall 0\leq t_1<t_2\leq1$，$\tilde{K}(t_1,t_2)\triangleq\{\int_{t_1}^{t_2} f dt : f\in\tilde{K}\}\subset\mathbb{R}$ 相对紧（由 $\int_0^1|f|\leq M$）

(ii) $\sup_{f\in\tilde{K}}\int_0^{1-h}|f(t+h)-f(t)|^p dt\leq\tilde{c}\tilde{\delta}^\alpha \xrightarrow{\tilde{\delta}\to0} 0$

$\therefore \sup_{n\geq1} P_{X_n}(\overline{\tilde{K}}^c) \leq \sup_{n\geq1} P_{X_n}(\tilde{K}^c) \leq \sup_{n\geq1}\mathbb{P}(\int_0^1|X_n(t)|dt>\tilde{M}) + \sup_{n\geq1}\mathbb{P}(\sup_{h\in(0,\tilde{\delta})}\int_0^{1-h}|X_n(t+h)-X_n(t)|^p dt>\tilde{c}\tilde{\delta}^{\tilde{\alpha}})$

（取 $\tilde{\delta}=\delta$，$\tilde{\alpha}=\alpha$，$\tilde{c}=C\tilde{M}$ 即可.）$\leq \frac{1}{\tilde{M}}\sup_{n\geq1}E\int_0^1|X_n|dt + \frac{1}{\tilde{c}\tilde{\delta}^{\tilde{\alpha}}}\sup_{n\geq1}E\left[\sup_{h\in(0,\tilde{\delta})}\int_0^{1-h}|X_n(t+h)-X_n(t)|^p dt\right]$. #

*Thm (Ascoli) 设 E: Banach，$B\subset C([0,1]:E)$，则

B 相对紧（$\bar{B}$ 紧）$\Leftrightarrow$ (i) $\forall t\in[0,1]$，$B(t)=\{f(t):f\in B\}\subset E$ 相对紧

(ii) B 等度连续，i.e. $\forall\varepsilon>0,\exists\delta,\forall|t-s|<\delta$，有 $\sup_{f\in B}|f(t)-f(s)|_E<\varepsilon$

*Thm (Arzela-Ascoli) 设 $B\subset C([0,1]:\mathbb{R})$，若

(i) B 逐点有界：$\forall t$，$B(t)$ 有界

(ii) B 等度连续：$\forall\varepsilon>0,\exists\delta>0$, s.t. $\forall|t-s|<\delta$，有 $\sup_{f\in B}|f(t)-f(s)|<\varepsilon$

则 B 相对紧.

*Thm. 设 E: Banach，$B\subset L^p([0,1]:E)$，则

B 相对紧 $\Leftrightarrow$ (i) $\forall 0\leq t_1<t_2\leq1$，$B(t_1,t_2)\triangleq\{\int_{t_1}^{t_2}f(t)dt : f\in B\}\subset E$ 相对紧

(ii) $\lim_{h\to0}\sup_{f\in B}\int_0^{1-h}|f(t+h)-f(t)|_E^p dt = 0$.

Thm (Helly selection) 设 d.f. $\{F_n\}_{n\geq1}$，则 $\exists$ subseq $\{F_{n_k}\}_{k\geq1}$, s.t.

$F_{n_k}\Rightarrow F$ (i.e., $\forall x\in C_F$, $F_{n_k}(x)\to F(x)$, $k\to\infty$).

其中 $F:\mathbb{R}\to[0,1]$ 单增右连续（若 $F(-\infty)=0$, $F(+\infty)=1$ 才成为 d.f.）

（称为 defective 分布函数，此时称 "$F_{n_k}\Rightarrow F$: vague convergence"）

c.e.g. $\forall n\in\mathbb{N}$，$\delta_n\in\mathcal{P}(\mathbb{R})\Rightarrow\delta_\infty$

其中 $F_n(x)=\delta_n((-\infty,x])=1_{[n,+\infty)}(x)$，$F(x)=\delta_\infty((-\infty,x])=0,\ \forall x\in\mathbb{R}$.

原因：$\forall K\subset\mathbb{R}$ 紧集，$\exists n$, s.t. $\delta_n(K)=0 \Rightarrow \{\delta_n\}$ 不一致胎紧

e.g. $\{\delta_{\frac{1}{n}}\}\Rightarrow\delta_0$，且 $\{\delta_{\frac{1}{n}}\}$ 一致胎紧

$F_{\frac{1}{n}}(x)=1_{[\frac{1}{n},\infty)}(x)$　$F_0(x)=1_{[0,\infty)}(x)$.

Thm (Helly cor.) 设 $\{\mu_n\}\subset\mathcal{P}(\mathbb{R})$ 且一致胎紧. $F_n(x)=\mu_n((-\infty,x])$.

则 $\exists$ subseq $\{F_{n_k}\}$, s.t. $F_n\Rightarrow F$ (i.e. $\forall x\in C_F$, $F_n(x)\to F(x)$.)

其中 F 为 d.f.

Pf of Helly's cor: 由 Helly，只须证 $F(-\infty)=0$，$F(+\infty)=1$.

由一致胎紧，$\forall \varepsilon>0$，$\exists M(\varepsilon)>0$，s.t. $\mu_n([-M,M]^c)\le\varepsilon$，$\forall n$

$\because \mu_n([-M,M]^c)=1-F_n(M)+F_n((-M)-)$

① $1-F_{n_k}(M)\le\varepsilon$，$\forall k$.

$\therefore \forall x>M$，$x\in C_F$，$F_{n_k}(x)\ge F_{n_k}(M)\ge 1-\varepsilon$

$\xRightarrow{F_{n_k}(x)\to F(x)} F(x)\ge 1-\varepsilon \Rightarrow F(+\infty)\ge 1-\varepsilon \xRightarrow{F\text{值域}[0,1]} F(+\infty)=1$

类似，② $F_{n_k}((-M)-)\le\varepsilon$，$\forall k$. $\Rightarrow F(-\infty)=0$.

Pf of prokhorov 良定：

step 1. $0\le d_p(\mu,\nu)\le 1$

step 2. $d_p(\mu,\nu)=d_p(\nu,\mu)$

step 3. $d_p(\mu,\nu)=0 \Leftrightarrow \mu=\nu$

($\Rightarrow$) $\exists \underset{\varepsilon_n\in C_{\mu,\nu}}{\varepsilon_n\downarrow 0}$，s.t. $\forall B\in\mathcal{B}_S$，$\mu(B)\le\nu(B^{\varepsilon_n})+\varepsilon_n$，$\nu(B)\le\mu(B^{\varepsilon_n})+\varepsilon_n$.

令 $n\to\infty$，由概率测度连续性，$\mu(B)\le\nu(\overline{B})$，$\nu(B)\le\mu(\overline{B})$

当 B 为闭集时，$\mu(B)=\nu(B)$.

由 π-λ 定理，$\mu(B)=\nu(B)$，$\forall B\in\mathcal{B}_S \Rightarrow \mu=\nu$.

Step 4. ($\triangle$ ineq). $\forall \mu,\nu,m\in\mathcal{P}(S)$，$d_p(\mu,\nu)\le d_p(\mu,m)+d_p(m,\nu)$

$\forall \varepsilon_1\in C_{\mu,m}$，$\varepsilon_2\in C_{m,\nu}$，$\forall B\in\mathcal{B}_S$

$$\begin{cases}\mu(B)\le m(B^{\varepsilon_1})+\varepsilon_1\\ m(B)\le\mu(B^{\varepsilon_1})+\varepsilon_1\end{cases}\qquad \begin{cases}\nu(B)\le m(B^{\varepsilon_2})+\varepsilon_2\\ m(B)\le\nu(B^{\varepsilon_2})+\varepsilon_2\end{cases}$$

$$\xRightarrow{B^{\varepsilon_1}\in\mathcal{B}_S}\begin{cases}\mu(B)\le\nu((B^{\varepsilon_1})^{\varepsilon_2})+\varepsilon_1+\varepsilon_2 \overset{(B^{\varepsilon_1})^{\varepsilon_2}\subset B^{\varepsilon_1+\varepsilon_2}}{\le} \nu(B^{\varepsilon_1+\varepsilon_2})+\varepsilon_1+\varepsilon_2\\ \nu(B)\le\mu((B^{\varepsilon_2})^{\varepsilon_1})+\varepsilon_1+\varepsilon_2\le\mu(B^{\varepsilon_1+\varepsilon_2})+\varepsilon_1+\varepsilon_2\end{cases}$$

$\Rightarrow d_p(\mu,\nu)\le\varepsilon_1+\varepsilon_2$

Rmk. 由 Helly selection 定理，设 $\mathcal{M}\subset\mathcal{P}(\mathbb{R})$，则

$\mathcal{M}$ 一致胎紧 $\Leftrightarrow$ $\mathcal{M}$ 相对紧（弱拓扑）

§4.5 prokhorov 定理

Def (prokhorov 度量) 设 $\mu,\nu\in\mathcal{P}(S)$，定义 $d_p(\mu,\nu)\triangleq\inf\{\varepsilon>0: \mu(B)\le\nu(B^{\varepsilon})+\varepsilon, \nu(B)\le\mu(B^{\varepsilon})+\varepsilon, \forall B\in\mathcal{B}_S\}$. ($\{\cdots\}\triangleq C_{\mu,\nu}$)

其中 (S,d) 度量空间，$B^{\varepsilon}\triangleq\{x\in S: d(x,B)\le\varepsilon\}$.

称 d_p 为 prokhorov 度量.

Thm (良定) 若 (S,d) 为度量空间，则 $(\mathcal{P}(S),d_p)$ 是度量空间.

Rmk. 1° (S,d) 是 Polish 空间 $\Rightarrow$ $(\mathcal{P}(S),d_p)$ 是 Polish 空间.

2° $d_p(\mu_n,\mu)\to 0 \Leftrightarrow \mu_n\Rightarrow\mu$.

($\Rightarrow$) $\exists \varepsilon_n\downarrow 0$，$\varepsilon_n\in C_{\mu_n,\mu}$，$\forall B\in\mathcal{B}_S$

$$\begin{cases}\mu_n(B)\le\mu(B^{\varepsilon_n})+\varepsilon_n\\ \mu(B)\le\mu_n(B^{\varepsilon_n})+\varepsilon_n.\end{cases}$$

令 $n\to\infty$，$\varlimsup_{n\to\infty}\mu_n(B)\le\mu(\overline{B})$，取 B 闭，$\varlimsup_{n\to\infty}\mu_n(B)\le\mu(B)$

由 Portmanteau，$\mu_n\Rightarrow\mu$. #

E.g. Warsserstein 度量 (S,d) 度量空间，$\forall p\ge 1$，定义

$\mathcal{P}_p(S)\triangleq\{\mu\in\mathcal{P}(S): \int_S d(x_0,y)^p\mu(dy)<\infty, \exists x_0\in S\}$ 为 Warsserstein 空间

Rmk. $\mathcal{P}_p(S)=\{\mu\in\mathcal{P}(S): \int_S d(x,y)^p\mu(dy)<\infty, \forall x\in S\}$.

∵ $d(x,y)^p\le 2^{p-1}(d(x,x_0)^p+d(x_0,y)^p)$

Def. $\forall\mu,\nu\in\mathcal{P}_p(S)$，$W_p(\mu,\nu)\triangleq\left\{\inf_{\pi\in\Pi(\mu,\nu)}\int_{S\times S}d(x,y)^p\pi(dx,dy)\right\}^{\frac{1}{p}}$ 为 p-阶 Warsserstein 度量.

其中 $\Pi(\mu,\nu)\triangleq\{\pi\in\mathcal{P}(S\times S): \pi(A\times S)=\mu(A), \pi(S\times B)=\nu(B), \forall A,B\in\mathcal{B}_S\}$

Rmk. 1° 称 π 为 transportation plan

2° W_p 在 $\mathcal{P}_p(S)$ 中确是一个度量.

∵ 有限性：$W_p^p(\mu,\nu)\underset{\mu\times\nu\in\Pi(\mu,\nu)}{\le}\int_{S\times S}d(x,y)^p\mu(dx)\nu(dy)\le 2^{p-1}\int_S d(x,x_0)^p\mu(dx)+2^{p-1}\int_S d(x_0,y)^p\nu(dy)<\infty$.

Eg. ($p=2$，平方-Warsserstein 度量) 设 $(S,\|\cdot\|)$ is Banach，$\mu\in\mathcal{P}_2(S)$，$a\in S$.

$$W_2^2(\mu,\delta_a)=\inf_{\pi\in\Pi(\mu,\delta_a)}\int_{S\times S}\|x-y\|^2\pi(dx,dy)$$
$$=\inf_{(X,Y)\in\widetilde{\Pi}(\mu,\delta_a)}E\|X-Y\|^2$$
$$=\inf_{(X,a)\in\widetilde{\Pi}(\mu,\delta_a)}E\|X-a\|^2\quad \because \widetilde{\Pi}(\mu,\delta_a)=\{(X,Y):P_X=\mu,P_Y=\delta_a\}=\{(X,a):P_X=\mu\}.$$
$$=\int_S\|x-a\|^2\mu(dx).$$

$\therefore \inf_{a\in S}W_2^2(\mu,\delta_a)=\int_S\|x-\int_S x\mu(dx)\|^2\mu(dx)$ 为方差.

($p=1$，Kantorowich-Rubinstein 度量) 设 (S,d) is Polish，$\mu,\nu\in\mathcal{P}_1(S)$

K-R 对偶公式. $W_1(\mu,\nu)=\inf_{\pi\in\Pi(\mu,\nu)}\int_{S\times S}d(x,y)\pi(dx,dy)$
$$=\sup_{g\in C^d_{Lip,1}(S)}\int_S g(x)(\mu(dx)-\nu(dx)),$$

其中 $C^d_{Lip,1}(S)\triangleq\{g:S\to\mathbb{R}\mid \|g\|_{Lip}=\sup_{x\neq y\in S}\frac{|g(x)-g(y)|}{d(x,y)}\le 1\}$

Pf of 全变差公式：取 $d(x,y)=\mathbb{1}(x\neq y)$ 为度量，则由 K-R 对偶.

$$W_1(\mu,\nu)=\inf_{\pi\in\Pi(\mu,\nu)}\pi(\{x\neq y\})$$
$$\|$$
$$\sup_{g\in C^d_{Lip,1}(S)}\int_S g(x)(\mu(dx)-\nu(dx))$$
$$\|\ \text{CLAIM}$$
$$\sup_{g\in\{g:S\to[0,1]\text{可测}\}}\int_S g(x)(\mu-\nu)(dx)=\int_S 1\cdot(\mu-\nu)^+(dx)=(\mu-\nu)^+(S)=\tfrac12|\mu-\nu|_{TV}$$

下证 CLAIM：1° $\{g:S\to[0,1]\text{可测}\}\subset C^d_{Lip,1}(S)=\{g:\sup_{x\neq y}|g(x)-g(y)|\le1\}$.

2° $\forall g\in C^d_{Lip,1}(S)$，取 $\tilde g=g-\inf g$，则 $\tilde g\in\{g:S\to[0,1]\text{可测}\}$

且 $\int_S\tilde g\,d(\mu-\nu)=\int_S(g-\inf g)\,d(\mu-\nu)=\int_S g\,d(\mu-\nu)$. #.

Ex. (协方差不等式) 设 (S,d) is polish，$f:S\to\mathbb{R}_+$，$\int_S fd\mu=1$，其中 $\mu\in\mathcal{P}_1(S)$

则 $(\int_S fd\mu)(\int_S gd\mu)-(\int_S fg\,d\mu)\le\|g\|_{Lip}W_1(\mu,f\mu)$.

$$\|$$
$$E[f(X)]E[g(Y)]-E[f(X)g(Y)]$$

其中 $g\in Lip$. $(f\mu)(B)\triangleq\int_B fd\mu$，$\forall B\in\mathcal{B}_S$

(Hint：不妨 $\|g\|_{Lip}=1$.)

(r.v. 表达)

3° 定义 $\widetilde{\Pi}(\mu,\nu)\triangleq\{(X,Y)$ 为 $(\Omega,\mathcal{F},\mathbb{P})$ 下 $S\times S$ 值 r.v.，$P_X=\mu,P_Y=\nu\}$.

称 $(X,Y)\in\widetilde{\Pi}(\mu,\nu)$ 为 μ,ν 的一个耦合 (coupling).

$$W_p(\mu,\nu)=\{\inf_{\pi\in\Pi(\mu,\nu)}\int_{S\times S}d(x,y)^p\pi(dx,dy)\}^{\frac1p}$$
$$=\{\inf_{(X,Y)\in\widetilde{\Pi}(\mu,\nu)}\int_{S\times S}d(x,y)^pP_{(X,Y)}(dx,dy)\}^{\frac1p}$$
$$=\{\inf_{(X,Y)\in\widetilde{\Pi}(\mu,\nu)}E[d(X,Y)^p]\}^{\frac1p}$$

(Recall：θ 符号测度，Jordan 分解 $\theta=\theta^+-\theta^-$.

全变差测度 $|\theta|\triangleq\theta^++\theta^-$，全变差 $|\theta|_{TV}\triangleq|\theta|(S)$.)

Thm. (全变差公式) 设 S is polish，$\mu,\nu\in\mathcal{P}_1(S)$，则

$$\inf_{\pi\in\Pi(\mu,\nu)}\pi(\{x\neq y\})=\tfrac12|\mu-\nu|_{TV}\left(=(\mu-\nu)^+(S)=(\mu-\nu)^-(S)\right)$$

Thm (Prokhov 定理) 设 (S,d) is polish，$M\subset\mathcal{P}(S)$，则

M 相对紧 $\Longleftrightarrow$ M 一致胎紧

[弱拓扑下，即 in $d_p(\cdot,\cdot)$ ($\Leftrightarrow$ 弱收敛)]

e.g. $S=C([0,T]:\mathbb{R}^m)$ 或 $L^p([0,T]:\mathbb{R}^m)$ 或 $D([0,T]:\mathbb{R}^m)$ (Skorohod 度量).

(另一方向难).

Pf：($\Rightarrow$) CLAIM：设 $\{U_k\}$ 开覆盖 S，$S=\bigcup_{k=1}^\infty U_k$

则 $\forall\varepsilon>0$，$\exists n$，s.t. $\forall\mu\in M$，有 $\mu(\bigcup_{k=1}^n U_k)>1-\varepsilon$.

证明：(反证) $\exists\varepsilon>0$，$\forall n$，$\exists\mu_n\in M$，s.t. $\mu_n(\bigcup_{k=1}^n U_k)\le1-\varepsilon$.

由 M 相对紧，$\exists\mu$ 概率测度，s.t. $\mu_{n_k}\Rightarrow\mu$

由 Portmanteau $\mu(\bigcup_{k=1}^n U_k)\le\varliminf_{k\to\infty}\mu_{n_k}(\bigcup_{k=1}^n U_k)$
$$\le\varliminf\mu_{n_k}(\bigcup_{k=1}^{n_k}U_k)\le1-\varepsilon$$

令 $n\to\infty$，$\mu(S)\le1-\varepsilon$ 与 $\mu(S)=1$ 矛盾！得证！

S 可分 $\Rightarrow D=\{a_1,a_2,\cdots\}$ 为 S 的可数稠密子集

$\Rightarrow\forall m\ge1$，$S=\bigcup_{k=1}^\infty B_{\frac1m}(a_k)$

$\overset{\text{由CLAIM}}{\Rightarrow}\exists n_m\ge1$，s.t. $\mu(\bigcup_{k=1}^{n_m}B_{\frac1m}(a_k))>1-\frac{1}{2^m}\varepsilon$，$\forall\mu\in M$.

取闭集 $K\triangleq\bigcap_{m\ge1}\bigcup_{k=1}^{n_m}\overline{B_{\frac1m}}(a_k)$，下证 K 紧 $\underset{S\text{完备}}{\Longleftrightarrow}$ K 完全有界.

$\because\forall\delta>0$，$\exists\frac1m<\delta$，$K\subset\bigcup_{k=1}^{n_m}\overline{B_{\frac1m}}(a_k)\subset\bigcup_{k=1}^{n_m}B_\delta(a_k)\Rightarrow K$ 完全有界.

另外，$\sup_{\mu\in M}\mu(K^c)\le\sup_{\mu\in M}\sum_{m\ge1}\mu([\bigcup_{k=1}^{n_m}\overline{B_{\frac1m}}(a_k)]^c)\le\sum_{m\ge1}\frac1{2^m}\varepsilon=\varepsilon$.

$\therefore M$ 一致胎紧 #.

Ex. (S,d) polish, $\{\mu_n\}\subset \mathcal{P}_1(S)$ (Warsenstein度量). 若 $\{\mu_n\}$ Cauchy in W_1

则 $\{\mu_n\}$ 一致胎紧.

Pf: $W_1(\mu_k,\mu_l)\to 0 \ (k,l\to\infty)$

$\Rightarrow \forall \varepsilon>0, \exists N$, s.t. $\forall k\geq N$, $W_1(\mu_k,\mu_N)\leq \varepsilon^2$

$\Rightarrow \forall k\geq 1, \exists j\in\{1,\cdots,N\}$ s.t. $W_1(\mu_k,\mu_j)\leq \varepsilon^2$

由 (S,d) polish, $\{\mu_1,\cdots,\mu_N\}$ 一致胎紧.

$\Rightarrow \exists K=K_\varepsilon\subset S$ 紧, s.t. $\mu_j(K^c)\leq\varepsilon$, $\forall j=1,\cdots,N$

另外, K 紧 $\Rightarrow$ 完全有界 $\Rightarrow \exists m, \{x_i\}_{i=1}^m\subset S$, s.t. $K\subset \bigcup_{i=1}^m B_\varepsilon(x_i)\triangleq U$

$\subset U^\varepsilon \subset T \triangleq \bigcup_{i=1}^m B_{2\varepsilon}(x_i)$

定义 $g_\varepsilon(x)\triangleq \left\{1-\frac{d(x,U)}{\varepsilon}\right\}^+$, $x\in S$,

则 g_ε 为 ε^{-1}-lip 函数. $\Rightarrow \varepsilon g_\varepsilon \in C^d_{Lip,1}(S)$.

由 $1_U\leq g_\varepsilon\leq 1_{U^\varepsilon}$

$\mu_k(U^\varepsilon)\geq \int g_\varepsilon d\mu_k=\int g_\varepsilon d\mu_j+\int g_\varepsilon(d\mu_k-d\mu_j)$

$\because W_1(\mu_k,\mu_j)\overset{K\text{-R对偶}}{=}\sup_{g\in C^d_{Lip,1}(S)}\int g\, d(\mu_j-\mu_k)\geq \int \varepsilon g_\varepsilon\, d(\mu_j-\mu_k)$

$\therefore \int g_\varepsilon(d\mu_k-d\mu_j)\geq -\frac{1}{\varepsilon}W_1(\mu_k,\mu_j)\geq -\varepsilon$.

$\therefore \mu_k(U^\varepsilon)\geq \mu_j(U)-\varepsilon\geq 1-2\varepsilon$.

$\therefore \mu_k(T)\geq 1-2\varepsilon$.

$\Rightarrow \forall l\geq 1, \exists m(l)\geq 1$, s.t. $\forall k\geq 1$, 有 $\mu_k\left(\bigcup_{i=1}^{m(l)} B_{2^{-l}\varepsilon}(x_i)\right)>1-\frac{1}{2^l}\varepsilon$

取闭集 $\bigcap_{n\geq 1}\bigcup_{i=1}^{m(l)}\overline{B_{2^{-l}\varepsilon}(x_i)}\triangleq C$, 完全有界: $\forall\delta>0, \exists 2^{-l}\varepsilon<\delta$, $C\subset\bigcup_{i=1}^{m(l)}\overline{B_{2^{-l}\varepsilon}(x_i)}$

$\subset \bigcup_{i=1}^{m(l)} B_\delta(x_i)$

再 check $\mu_k(C^c)\leq\varepsilon$ 即可.

#

Rmk. W_1 换成 W_p, 由Hölder, $W_p\geq W_1$, 上述结论也对

§4.6 一致可积. (Uniformly Integrable, U.I.)

一致可积 $\Rightarrow$ 一致 L^1 有界 $\Rightarrow$ 一致胎紧.

Recall: D.C.T.

$$\left.\begin{array}{l} X_n\xrightarrow{a.e.}X \\ |X_n|\overset{a.e.}{\leq} Y\in L^1\end{array}\right\}\Rightarrow EX_n\to EX,\ E|X_n|\to E|X|$$

$$\left.\begin{array}{l} X_n\xrightarrow{\mathbb{P}}X \\ |X_n|\overset{a.e.}{\leq} Y\in L^p\end{array}\right\}\Rightarrow X_n\xrightarrow{L^p}X\overset{p=1\text{时}}{\Rightarrow} E|X_n|\to E|X|.$$

CLAIM:
$$\left.\begin{array}{l} X_n\xrightarrow{\mathbb{P}}X \\ \{X_n\}\ U.I.\end{array}\right\}\Rightarrow X_n\xrightarrow{L^1}X\Rightarrow E|X_n|\to E|X|$$

Def. (U.I.) 设 $(\Omega,\mathcal{F},\mathbb{P})$ 下实值 r.v.s $\{X_n\}_{n\geq 1}$.

若 $\lim_{M\to\infty}\sup_{n\geq 1}E\left[|X_n|1_{|X_n|>M}\right]=0$, 称 $\{X_n\}$ 一致可积, 记为U.I

Rmk. 1° $\{X_\alpha\}$ is U.I. $\Leftrightarrow \lim_{M\to\infty}\sup_{\alpha\in I}E\left[|X_\alpha|1_{|X_\alpha|>M}\right]=0$.

2° $|X_n|\overset{a.e.}{\leq} Y\in L^1(\mathcal{F})\Rightarrow \{X_n\}$ is U.I.

$E\left[|X_n|1_{|X_n|>M}\right]\leq E\left[Y\cdot 1_{Y>M}\right]$,

而 $Y\in L^1$, $\mathbb{P}(Y>M)\to 0\Rightarrow E\left[Y\cdot 1_{Y>M}\right]\to 0$

$\therefore \{X_n\}$ is U.I.

3° $\{X_t\}_{t\geq 0}$. 若 $\sup_{t\geq 0}|X_t|\in L^1(\mathcal{F})$, 则 $\{X_t\}_{t\geq 0}$ is U.I.

4° $\{X_1,\cdots,X_N\}$ 可积 $\Rightarrow$ U.I. 取 $Y=\sum_{i=1}^N|X_i|$

Lem. (U.I. $\Rightarrow$ 一致 L^1 有界). $\{X_n\}$ is U.I. $\Rightarrow \{X_n\}$ is 一致 L^1 有界, 即 $\sup_{n\geq 1}E|X_n|<\infty$

从而 $\{X_n\}$ 一致胎紧

Pf: $\forall M>0$, $\sup_{n\geq 1}E|X_n|\leq\sup_{n\geq 1}E\left[|X_n|1_{|X_n|>M}\right]+\sup_{n\geq 1}E\left[|X_n|1_{|X_n|\leq M}\right]$

$\leq\sup_{n\geq 1}E\left[|X_n|1_{|X_n|>M}\right]+M<\infty$. #

U.I.判别法:

Thm. (U.I $\Leftrightarrow$ 一致 L^1 有界 + a.c.) $\{X_n\}$ is UI $\Leftrightarrow$ ① $\{X_n\}$ 一致 L^1 有界 ($\sup E|X_n|<\infty$)

② $\forall\varepsilon>0, \exists\delta$, s.t. $\forall \mathbb{P}(A)<\delta$, 有 $E|X_n|1_A<\varepsilon$, $\forall n$

<u>Pf</u> of "U.I $\Leftrightarrow$ 一致L^1有界 + a.c."

($\Rightarrow$) $E|X_n| 1_A = E|X_n| 1_{A\cap\{|X_n|>M\}} + E|X_n| 1_{A\cap\{|X_n|\le M\}}$

$\le E|X_n| 1_{\{|X_n|>M\}} + M\cdot \mathbb{P}(A)$.

$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2}$ (取M充分大, $\delta = \frac{\varepsilon}{2M}$).

($\Leftarrow$) ① $\Rightarrow \mathbb{P}(|X_n|>M) \le \frac{1}{M}\sup E|X_n| < \delta$ (取M充分大).

$\Rightarrow E|X_n| 1_{\{|X_n|>M\}} < \varepsilon, \forall n\ge 1$. #

<u>Pf</u> of UITF 判别:

($\Rightarrow$) 记 $M = \sup_{\alpha\in I} E\varphi(|X_\alpha|) < \infty$

由 φ sublinear, $\exists$ 正数列 $\{C_n\}$, s.t. $\lim_{n\to\infty} C_n = \infty$, 且 $\varphi(x) \ge n\cdot M x$, $\forall x > C_n$

$\therefore \sup_{\alpha\in I} E[|X_\alpha| 1_{|X_\alpha|>C_n}] \le \frac{1}{nM}\sup_{\alpha\in I} E\varphi(|X_\alpha|) = \frac{1}{n}$ $\therefore \{X_\alpha\}$ is U.I.

($\Leftarrow$) 记 $f(K) = \sup_{\alpha\in I} E[|X_\alpha| 1_{\{|X_\alpha|>K\}}]$, 则 $f(K)\downarrow 0$

$\therefore \exists g: \mathbb{R}^+ \to \mathbb{R}^+$, s.t.

$\int_0^\infty g(y)dy = \infty$ 且 $\int_0^\infty g(y) f(y) dy < \infty$

定义 $\varphi: \mathbb{R}^+ \to \mathbb{R}^+$

$x \mapsto x\int_0^x g(y)dy$, 则 φ 为 UITF.

Check: $\sup_{\alpha\in I} E\varphi(|X_\alpha|) < \infty$.

$= \sup_{\alpha\in I} E[|X_\alpha| \int_0^{|X_\alpha|} g(y)dy]$

$= \sup_{\alpha\in I} E[|X_\alpha| \int_0^\infty g(y) 1_{y<|X_\alpha|} dy]$

$\overset{\text{Fubini}}{\le} \int_0^\infty \sup_{\alpha\in I} E[|X_\alpha| 1_{|X_\alpha|>y}] g(y) dy$

$= \int_0^\infty f(y) g(y) dy < \infty$. #

(续右) $X_n \xrightarrow{\mathbb{P}} X \Rightarrow |X_n| \xrightarrow{\mathbb{P}} |X| \Rightarrow |X_n| \xrightarrow{d} |X|$

$\underline{\lim} E|X_n| 1_{|X_n|\le M} \ge \lim E g_M(|X_n|) = E g_M(|X|) \ge E|X| 1_{|X|\le M-1}$

$\Rightarrow \overline{\lim} E|X_n| 1_{|X_n|>M} \le E|X| 1_{|X|>M-1} (\le E|X| < \infty)$.

$\exists M_0, \exists m\ge 1$, s.t. $\forall M\ge M_0$, $\sup_{n\ge m} E|X_n| 1_{|X_n|>M} < \varepsilon \Rightarrow \exists M_1 > M_0$, s.t. $\forall M \ge M_1$, $\sup_{n\ge 1}[E|X_n| 1_{|X_n|>M}] < \varepsilon$ #

• 一致可积性的判别 (de la vallée-Poussin 判别).

<u>引例</u>. $\{X_n\}$ 为一列 r.v.s, $\forall \varepsilon > 0$.

$\sup_{n\ge 1} E|X_n| 1_{\{|X_n|>M\}} \le \frac{1}{M^\varepsilon}\sup_{n\ge 1} E|X_n|^{1+\varepsilon}$.

函数 $\varphi(x) \triangleq x^{1+\varepsilon}$, $x>0$, $\underbrace{\lim_{x\to\infty}\frac{\varphi(x)}{x} = +\infty}$ 满足该条件, 称为一致可积 test function (superlinear function)

则 $\sup_{n\ge 1} E\varphi(|X_n|) < \infty \Rightarrow \{X_n\}$ is U.I.

<u>Def</u>. (一致可积测试函数, U.I.T.F.) 设 $\varphi: \mathbb{R}_+ \to \mathbb{R}_+$ 为满足 $\lim_{x\to\infty}\frac{\varphi(x)}{x} = +\infty$ 的函数, 称 φ 为一个 UITF.

<u>Thm</u>. (UITF 判别 UI). 设 $\{X_\alpha\}_{\alpha\in I}$ 为一族 r.v.s, 则

$\exists$ U.I.T.F. φ, s.t. $\sup_{\alpha\in I} E[\varphi(|X_\alpha|)] < \infty$

$\Longleftrightarrow \{X_\alpha\}$ is U.I.

<u>Thm</u> (Vitali 收敛定理). 设 $\{X_n\} \subset L^p(\mathcal{F})$, $p\ge 1$ 且 $X_n \xrightarrow{\mathbb{P}} X$, 则下面条件等价.

① $\{X_n^p\}$ is U.I.

② $X_n \xrightarrow{L^p} X$

③ $E|X_n|^p \to E|X|^p$.

<u>Pf</u>: 只证 $p=1$. $\left.\begin{array}{l} X_n \in L^1(\mathcal{F}) \\ X_n \xrightarrow{\mathbb{P}} X \end{array}\right\} \overset{\text{Fatou}}{\Longrightarrow} X \in L^1(\mathcal{F})$

① $\Rightarrow$ ②. $E|X_n - X| = E[|X_n - X| 1_{|X_n-X|>\delta}] + E[|X_n-X| 1_{|X_n-X|\le\delta}]$

$\le E[|X_n - X| 1_{|X_n-X|>\delta}] + \delta$.

$\forall \varepsilon>0, \exists n_0$, s.t. $\forall n\ge n_0$, $\mathbb{P}(|X_n - X|>\delta) < \varepsilon$.

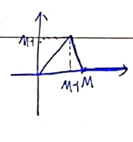
$\{X_n\}$ is U.I $\Rightarrow \{|X_n|\}$ is U.I. $\Rightarrow \{|X_n| + |X|\}$ is U.I.

$\Rightarrow \overline{\lim} E|X_n - X| \le \delta$

由 δ 任意, $\overline{\lim} E|X_n - X| = 0 \Rightarrow X_n \xrightarrow{L^1} X$.

② $\Rightarrow$ ③ $E|X_n - X| \to 0 \Rightarrow E|X_n| \to E|X|$.

③ $\Rightarrow$ ① 定义 $g_M(x) = \begin{cases} x, & 0\le x< M-1 \\ 0, & x\ge M \\ \text{插值}, & M-1\le x< M \end{cases}$



则 $x\cdot 1_{\{x\le M-1\}} \le g_M \le x 1_{\{x\le M\}}$ $\in C_b$.

助教版本：$(S,\mathcal{S})$ is polish，$\forall A\in\mathcal{S}$，$\exists F$闭，G开，$F\subset A\subset G$.

$\forall \mathbb{P}\in\mathcal{P}(S)$，$\mathbb{P}(\ /F)<\varepsilon$.

Pf：① 若 A 闭，$F=A$．$G^{\delta}=\{x\in S \mid d(x,A)<\delta\}$ $\bigcap_{n=1}^{\infty}G^{\delta_n}=A=F$ $(\delta_n\to 0)$

$\exists\, n$ 足够大，s.t. $\mathbb{P}(\bigcap_{n=1}^{N}G^{\delta_n}/F)<\varepsilon$.

② $\mathcal{F}\triangleq\{A\in\mathcal{S}:\forall\varepsilon>0,\exists F闭\subset A\subset G开,\mathbb{P}(G/F)<\varepsilon\}$，下证为 σ-alg.

(i) $\Omega\in\mathcal{F}$.

(ii) 若 $A\in\mathcal{F}$，则 $A^c\in\mathcal{F}$．$\because$ $G\overset{闭}{\subset}A\overset{开}{\subset}F\Rightarrow G^c\overset{开}{\supset}A^c\overset{闭}{\supset}F^c$.

(iii) 若 $A_n\in\mathcal{F}$，则 $\bigcup A_n\in\mathcal{F}$ $\because$ $\mathbb{P}(G_n\setminus F_n)<\frac{\varepsilon}{2^{n+1}}$

$\bigcup G_n$ 开 $\supset\bigcup A_n$

$\exists n_0$, s.t. $\mathbb{P}(\bigcup_{n=1}^{\infty}F_n/\bigcup_{n=1}^{n_0}F_n)<\frac{\varepsilon}{2}$

$\mathbb{P}(\bigcup_{n=1}^{\infty}G_n\setminus\bigcup_{n=1}^{n_0}F_n)=\mathbb{P}(\bigcup_{n=1}^{\infty}G_n/\bigcup_{n=1}^{\infty}F_n)$

$+\mathbb{P}(\bigcup_{n=1}^{\infty}F_n/\bigcup_{n=1}^{n_0}F_n)<\varepsilon$

#

△ Ex. 设 S 为拓扑空间，$\mu\in\mathcal{P}(S)$．则 $\forall B\in\mathcal{B}_S$，有.

$\mu(B)=\inf_{B\subset U,U开}\mu(U)=\sup_{D\subset B,D闭}\mu(D)$．称 μ 正则.

($\forall$ Borel 概率测度都正则)

Hints：① $\mathcal{M}\triangleq\{B\in\mathcal{B}_S:\mu(B)=\inf_{B\subset U开}\mu(U)=\sup_{B\supset D闭}\mu(D)\}$ 为 σ-alg.

② $\{闭集D\}\subset\mathcal{M}$.

$\Rightarrow\mathcal{B}_S\subset\mathcal{M}\subset\mathcal{B}_S$．$\Rightarrow\mathcal{M}=\mathcal{B}_S$

考试重点：{ 逼近

π-λ类定理.