2.6 Measure-Probality

prop: any set Ω with an outer measure, (ie. a mapping from $P(\Omega)$ to $[0,+\infty)$ which is monotone and countably additive) can be equipped with a measure μ defined on an appropriate σ-field F of its subset $\Rightarrow$ obtain the measure space (Ω, F, μ)

prop2.20 P45 (Restriction)

$B \subseteq \mathbb{R}$ is lebesgue measurable with $m(B)>0$; give measure m on the lebesgue σ-field M on $\mathbb{R}$,

let $M_B := \{A \cap B : A \in M\}$; and for $C \in M_B$, let $m_B(C) := m(C)$

$\Rightarrow$ now we restrict $(\mathbb{R}, M, m)$ to (B, M_B, m_B)

$$\begin{cases} M_B := \{A \cap B : A \in M\} \text{ is a } \sigma\text{-algebra} \\ (B, M_B, m_B) \text{ is a complete measure space} \end{cases}$$

def2.6 P46 A probability space is a triple (Ω, F, P) where Ω is a set, F is a σ-field of subsets of Ω, P is a measure on F s.t. $P(\Omega)=1$; P is called "probability" 默认 $P(\emptyset)=0$

rmk: elements in F are called "events"

Ω中包含了所有可能情况,但并不是每一种都成为合法的"event",

事件之间必须可以取补,可数交,可数并;这在古典概型中解释为 $P(A)+P(A^c)=1$,

可数次的 $P(A\cup B)=P(A)+P(B)-P(A\cap B)$; $A^c, \cup A_n, \cap A_n \in F$, 故要在Ω中取σ代数

rmk (Restriction on probability space)

in prop2.20, we obtain (B, M_B, m_B), $P(A) = \frac{m_B(A)}{m(B)} = \frac{m(A)}{m(B)}$ 归一化

3.5 Measurable function

随机变量=可测函数,分布=推前测度(pushforward measure)

(Ω, F, P) is a probability space, then $X: \Omega \to \mathbb{R}$ is a random variable if $X^{-1}([a,+\infty))$ in F for $\forall a \in \mathbb{R}$

ie. $\{w \in \Omega : X(w) \geq a\} \in F$ $\forall a \in \mathbb{R}$, $\Rightarrow$ $X: \Omega \to \mathbb{R}$ is ⋯ ; X is a Borel measurable function

3.52 σ-field generated by random variables

$\mathcal{B} \subseteq \mathbb{R}^1$ is a Borel σ-algebra, then $X^{-1}(\mathcal{B}) := \{S \subseteq F : S = X^{-1}(B) \text{ for some } B \in \mathcal{B}\}$ is a Borel σ-algebra,

$X^{-1}(\mathcal{B}) \subseteq F$, denoted F_X 这可以用来做 F_X 上的测度空间的restriction

rmk: 经典概率中,随机变量X是随机实验的结果;但在测度论中,这些随机实验应该要是"合法概率"

def: Probability distribution

$P_X(A) = P(X\in A) = P(X^{-1}(A))$, where $X:\Omega\to\mathbb{R}$ is a random variable, $A\subseteq\mathbb{R}$ 注意P.P_X定义域不一样Ω和R

if $P(X\in A, Y\in B) = P_X(A)\cdot P_Y(B)$, $\forall A,B\subseteq\mathbb{R}$, then $\sigma(X)$ independent with $\sigma(Y)$, X independent with Y

独立性是σ-代数之间的关系

4.7 Integral

TH4.28 given random variable $X:\Omega\to\mathbb{R}$, $\int_\Omega g(X(w))\,dP(w) = \int_{\mathbb{R}} g(x)\,dP_X(x) \Rightarrow E[g(X)]$

Pf: g is Borel function, $g:\mathbb{R}\to\mathbb{R}$; g measurable 默认可积

- let Borel set $B\in\mathcal{B}(\mathbb{R})$, let $g=1_B$ then:

$$\int_\Omega 1_B(X(w))\,dP(w) = \int_\Omega 1_{X^{-1}(B)}(w)\,dP(w) = P(X^{-1}(B)) = P_X(B) = \int_{\mathbb{R}} 1_B(x)\,dP_X(x) \quad \cdots ①$$

$X^{-1}(B)$中每一点加权$P(w)$　　类似的

- let $g(x)=\sum_{i=1}^n a_i\cdot 1_{B_i}(x)$, $a_i\geq 0$.

$$\int_\Omega g(X(w))\,dP(w) = \int_\Omega \sum_{i=1}^n a_i\cdot 1_{B_i}(X(w))\cdot dP(w) = \sum_{i=1}^n a_i\int_\Omega 1_{X^{-1}(B_i)}(w)\,dP(w) \quad \Leftarrow 代入①$$

$$= \sum_{i=1}^n a_i\int_{\mathbb{R}} 1_{B_i}(x)\,dP_X(x)$$

$$= \int_{\mathbb{R}} \overbrace{\left(\sum_{i=1}^n a_i\cdot 1_{B_i}(x)\right)}^{g}\,dP_X(x)$$ 即 LHS=RHS $\cdots ②$

- $g:\mathbb{R}\to[0,+\infty)$ is Borel measurable, then $\exists$ simple functions $\{g_n\}_{n=1}^{+\infty}$ s.t. $g_n\geq 0$, $g_n\uparrow g$ pointwise

$\therefore g_n(X(w))\uparrow g(X(w))$ for $\forall w$, $g_n(x)\uparrow g(x)$ $\forall x$ $\Leftarrow$ 代入②

$\int_\Omega g(X(w))\,dP(w) = \int_\Omega \lim g_n(X(w))\,dP(w)$, here "$\int\lim \rightleftharpoons \lim\int$" by TH MCT

$\lim\int_\Omega g_n(X(w))\,dP(w) = \lim\int_{\mathbb{R}} g_n(x)\,dP_X(x)$

$\Rightarrow \int_\Omega g(X(w))\,dP(w) = \int_{\mathbb{R}} g(x)\,dP_X(x)$

- let $g=g^+-g^-$, $g^+:\mathbb{R}\to[0,+\infty)$. $g^-:\mathbb{R}\to[0,+\infty)$

g^+, g^- satisfies the equation, thus g satisfies $\cdots$ $\Rightarrow$ 这里要求g^+, g^- integrable, 不然"$\pm\infty$"加减无定义

用到的引理: TH (lebesgue approximation)

$(X,\mathcal{A})$ is a measurable space, $f:X\to[0,+\infty)$ measurable; then $\exists$ simple functions $\{\varphi_n\}_{n=1}^{+\infty}$,

s.t. $\varphi_n\geq 0$ $\forall n$; $\varphi_n(x)\uparrow f(x)$ $\forall x$

其中: $\varphi_n(x)=\sum_{i=1}^{N_n} a_{n,i}\cdot 1_{A_{n,i}}$; $a_{n,i}\geq 0$, $A_{n,i}\in\mathcal{A}$ 可测函数可以被单调,非负,可测、简单fn逐点逼近

TH (Beppo-Levi / Monotone convergence TH)

$(X,\mathcal{A},\mu)$ is a measure space, $\{f_n:X\to[0,+\infty)\}_{n=1}^{+\infty}$ is measurable, $f_n(x)\uparrow f(x)$ almost everywhere,

then $\lim\int_X f_n\,d\mu = \int_X \lim f_n\,d\mu$　可测函数可以被单调,非负,可测逼近, lim可交换(不必可积!)

def P106 measure P of the form: $A \mapsto P(A) = \int_A f\,dm$ with non-negative、integrable f is called "absolutely continuous"; f is called a density (of P w.r.t Lebesgue measure)

P is a probability measure, then: $\int f\,dm = 1$ 存在密度 $\Leftrightarrow$ P测度对lebesgue绝对连续

def (cumulative distribution function) P110:

$X: \Omega \to \mathbb{R}$ is a random variable, $(\Omega, \mathcal{F}, P)$ is probability space;

def $F_X(y) := P(\{w: X(w) \le y\}) = P_X((-\infty, y])$

prop4.30 (1): F_X is non-decreasing $\uparrow$. (2) $\lim_{x\to+\infty} F_X(x) = 1$, $\lim_{x\to-\infty} F_X(x) = 0$

(3): F_X is right continuous, $\lim_{y\to y_0^+} F_X(y) = F_X(y_0)$ 不一定左连续,不一定连续

THM4.31 function $F: \mathbb{R} \to [0,1]$ satisfies (1)-(3) in prop4.30, $\Rightarrow \exists$ random variable $X: [0,1] \to \mathbb{R}$ on the probability space $([0,1], \mathcal{B}, m_{[0,1]})$ s.t. $F = F_X$ is the cdf 有硬归一化,其它也可以的

pf: $w \in [0,1]$, let $X^+(w) := \inf\{x: F(x) > w\}$, $X^-(w) := \sup\{x: F(x) < w\}$

$X^+(w) \ge X^-(w)$ obv, 接下来证明 $F_{X^-} = F$, 和(3)性质有关

$\Leftarrow F_{X^-}(y) := P(\{w: X^-(w) \le y\}) = P(\{w: \sup\{x: F(x) < w\} \le y\}) = F(y)$

$\parallel$

这是个0开始的区间 $[0,a)$ 或 $[0,a]$, $a \le 1$, $P([0,a)) = P([0,a]) = a$

$\Leftarrow a = F(x)$ 可以取到!

$\Leftarrow X^-(w) \le y$ implies $w \le F(y)$, $w \le F(y)$ implies $X^-(w) \le y$

证一个方向:

$\sup\{x: F(x) < w\} \le y$, let $g(w) :=$ LHS, $g(w)\uparrow$

$\therefore$ if $w > F(y)$, LHS $> \sup\{x: F(x) < F(y)\} = y$ if the distribution is cts, Contradict!

$\therefore$ if $\exists\, x_1 < x_2$ s.t. $F(x_1) < w$, $F(x_2) \ge w$

$$\left.\begin{array}{l} \sup\{x: F(x) < w\} \le y \text{ implies } x_1 \le y \\ \text{if } w > F(y),\ F(x_2) \ge w > F(y),\ x_2 > y \end{array}\right\} \Rightarrow F \text{ not right-cts at } y.\ \text{contradict!}$$

$\Rightarrow X^-(w) \le y$ implies $w \le F(y)$

书上写的, $X^-(w) \le y$ then $F(X^-(w)) \le F(y)$; WTS: $w \le F(y)$, it suffices to show $w \le F(X^-(w))$

if $w > F(X^-(w))$, F is right-cts;

$\exists\, x_0$ s.t. $F(X^-(w)) < F(x_0) < w$; thus $x_0 \in \{x: F(x) < w\}$ contradict with "$X^-(w) = \sup\{x: F(x) < w\} \ge x_0$";

TH4.32 P112. P_X defined on $\mathbb{R}^n$ is absolutely continuous with density f_X; $g:\mathbb{R}^n\to\mathbb{R}'$ is integrable with respect to P_X, then: $\int_{\mathbb{R}^n} g(x)\,dP_X(x) = \int_{\mathbb{R}^n} f_X(x)\,g(x)\,dx$

即: $\forall A\in\mathcal{B}(\mathbb{R}^n)$, $P_X(A)=\int_A f_X(x)\,dx$ or, $dP_X(x)=f_X(x)\,dx$

pf: let $g=1_A \to$ let $g=\sum^{N} a_n\cdot 1_{A_n} \to$ let g be approximated $\to g=g^+-g^-$ 参照TH4.28方法

coro4.33 $\int_\Omega g(X)\,dP_X(x) = \int_\Omega f_X(x)\,g(x)\,dx$, for $\forall$ set $\Omega\in\mathcal{B}(\mathbb{R}^n)$

TH4.34 P113. $g:\mathbb{R}\to\mathbb{R}$ is increasing, differentiable (thus invertible), then: $f_{g(X)}(y)=f_X(g^{-1}(y))\cdot\frac{d}{dy}g^{-1}(y)$

pf: $F_{g(X)}(y)=P(g(X)\le y)=P(X\le g^{-1}(y))=F_X(g^{-1}(y))$

$dP_X(x)=f_X(x)\,dx \Rightarrow dF_X(x)=f_X(x)\,dx$

$\therefore f_{g(X)}(y)\,dy = f_X(g^{-1}(y))\,d(g^{-1}(y))$

if g is decreasing, $f_{g(X)}(y)\,dy = -f_X(g^{-1}(y))\,d(g^{-1}(y))$

$P(g(X)\le y)=P(X\ge g^{-1}(y)) = 1-F_X(g^{-1}(y))$

def4.5 P116 X is a random variable. $\varphi_X(t):=E(e^{itX})$ $t\in\mathbb{R}$; φ_X is the characteristic function of X

$\varphi_X(t)=\int e^{itx}\,dP_X(x) = \int e^{itx} f_X(x)\,dx$ in absolutely cts case

TH4.35, by def4.5 function φ_X satisfies:

(1) $\varphi_X(0)=1$, $|\varphi_X(t)|\le 1$ $\forall t\in\mathbb{R}'$

(2) $\varphi_{aX+b}(t)=E(e^{it(aX+b)})=\int e^{it(ax+b)}\cdot f_X(ax+b)\,dx = e^{itb}\cdot\varphi_X(at) \Rightarrow \varphi_{aX+b}(t)=e^{itb}\varphi_X(at)$

(2)' $\varphi_{aX+bY}(t)=\varphi_{aX}(t)\cdot\varphi_{bY}(t)=\varphi_X(at)\cdot\varphi_Y(bt) \Rightarrow \varphi_{aX+bY+c}(t)=e^{itc}\varphi_X(at)\cdot\varphi_Y(bt)$ if X,Y independent