第三章 条件数学期望

§1. 引例：设(X,Z)为$(\Omega,\mathcal{F},\mathbb{P})$下一个取值于$\{(a_i,b_j): i=1,\cdots,m, j=1,\cdots,n\}$的离散型r.v.

定义分布律 $p_{ij}=\mathbb{P}(X=a_i,Z=b_j)$, $i=1,\cdots,m$, $j=1,\cdots,n$

$$f(b_j)\triangleq E[X|Z=b_j]=\sum_{i=1}^m a_i\mathbb{P}(X=a_i|Z=b_j)=\sum_{i=1}^m a_i\frac{p_{ij}}{\mathbb{P}(Z=b_j)}=\frac{\sum_{i=1}^m a_ip_{ij}}{\sum_{i=1}^m p_{ij}}$$

$f(b_j)$是一个关于b_j的确定性函数

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

称$E[X|\sigma(Z)]$(或$E[X|Z]$)为r.v. X关于σ-代数$\sigma(Z)$的条件数学期望

问题：对于任意$X\in L^1(\mathcal{F})$，以及任意一个σ-代数$\mathcal{G}\subset\mathcal{F}$，如何得到$E[X|\mathcal{G}]$

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

另外，$\{G_j\}_{j=1}^n$为Ω的一个划分，故Z是一个简单r.v.，则

$\sigma(Z)=\sigma(\{G_1,\cdots,G_n\})=\{\bigcup_{j\in J}G_j: J\subset\{1,2,\cdots,n\}\}$ (J可取ϕ, $\bigcup_{j\in\phi}G_j=\phi$)

这样有 $\forall G\in\sigma(Z)$, $\exists J\subset\{1,\cdots,n\}$ 使 $G=\bigcup_{j\in J}G_j$

$$E[\underbrace{f(Z)}_{E[X|\sigma(Z)]}1_G]=\sum_{j\in J}E[f(Z)1_{G_j}]=\sum_{j\in J}E[f(b_j)1_{G_j}]$$

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

$$=\sum_{j\in J}\sum_{i=1}^m a_ip_{ij}$$

$$E[X1_G]=\sum_{j\in J}E[X1_{G_j}]=\sum_{j\in J}E[X1_{\{Z=b_j\}}]=\sum_{j\in J}\sum_{i=1}^m a_i\cdot 1\cdot\mathbb{P}(X=a_i,Z=b_j)$$

$$=\sum_{j\in J}\sum_{i=1}^m a_ip_{ij}$$

总结：$E[X|\sigma(Z)]=f(Z)$满足

(i) $E[X|\sigma(Z)]$是一个$\sigma(Z)$可测的 ($f(Z):(\Omega,\sigma(Z))\to(\mathbb{R},\mathcal{B}_{\mathbb{R}})$, $\forall B\in\mathcal{B}_{\mathbb{R}}$, $f^{-1}(B)\in\sigma(Z)$)

(ii) $\forall G\in\sigma(Z)$, $E[E[X|\sigma(Z)]1_G]=E[X1_G]$

<u>定理</u>(条件数学期望的存在唯一性)

设$X\in L^1(\mathcal{F})$以及$\mathcal{G}\subset\mathcal{F}$为任意$\sigma$-代数，则存在一个可积r.v. $Y\in\mathcal{G}$使

$$E[Y1_G]=E[X1_G],\quad\forall G\in\mathcal{G}$$

如果存在另一个可积r.v. $\hat{Y}\in\mathcal{G}$使上述结论成立，则$Y=\hat{Y}$ a.e.

注：以后记$E[X|\mathcal{G}]\triangleq Y$，称$E[X|\mathcal{G}]$为r.v. X关于σ-代数$\mathcal{G}$的条件数学期望

下面证存在唯一性.(道路漫长)

定义(相互奇异的测度)：设μ,ν为$(S,\mathcal{S})$上两个σ-有限测度

如果存在$G\in\mathcal{S}$，使$\mu(G)=\nu(G^c)=0$.

则称μ与ν相互奇异，记$\mu\perp\nu$

定理(测度的Leb分解)：设μ,ν为$(S,\mathcal{S})$上两个σ-有限测度，则

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

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

测度$\nu_s\triangleq\nu-f\mu\perp\mu$

进一步，若μ给定，则上面的Leb分解是唯一的

证：思路：找到一个非负可测f，使$\nu_s\triangleq\nu-f\mu\perp\mu$

为保证ν_s为测度，则$\forall B\in\mathcal{S}$，$\nu_s(B)=\nu(B)-f\mu(B)\geq 0$

故$f\in\mathcal{H}\triangleq\{f\text{非负可测}:\int_B f\,d\mu\leq\nu(B),\forall B\in\mathcal{S}\}$

下面假设μ,ν为$(S,\mathcal{S})$上有限测度

则$\mathcal{H}$有如下性质

(i) $h_n\in\mathcal{H}, n\geq 1, h_n\uparrow h\Rightarrow h\in\mathcal{H}$

(事实上，$h_n\in\mathcal{H}\Rightarrow\int_B h_n\,d\mu\leq\nu(B)\xrightarrow{MCT}\int_B h\,d\mu=\lim_{n\to\infty}\int_B h_n\,d\mu\leq\nu(B)\Rightarrow h\in\mathcal{H}$)

(ii) $h_1,h_2\in\mathcal{H}\Rightarrow h_1\vee h_2\in\mathcal{H}$

(事实上，$\forall B\in\mathcal{S}$，$\int_B h_1\vee h_2\,d\mu=\int_{B\cap\{h_1\leq h_2\}}h_2\,d\mu+\int_{B\cap\{h_1\leq h_2\}^c}h_1\,d\mu$

$\leq\nu(B\cap\{h_1\leq h_2\})+\nu(B\cap\{h_1\leq h_2\}^c)=\nu(B)$

故$h_1\vee h_2\in\mathcal{H}$)

定义$0\leq K\triangleq\sup_{h\in\mathcal{H}}\int_S h\,d\mu=\sup_{h\in\mathcal{H}}h\mu(S)\leq\nu(S)<+\infty$

$\exists\{h_n\}_{n\geq1}\subset\mathcal{H}$满足$\int_S h_n\,d\mu>K-\frac{1}{n}$

再定义$f_n\triangleq\max\{h_1,\cdots,h_n\}, n\geq 1$

则 f_n 非负可测且 $f_n\uparrow$ 且 $f_n\geq h_n, \forall n\geq 1$

由 $\mathcal{H}$ 的性质(ii)有 $f_n\in\mathcal{H}, \forall n\geq 1$，再由性质(i)得 $f\triangleq \lim_n f_n\in\mathcal{H}$

另外一方面 $\because f_n\geq h_n, \forall n\geq 1 \quad \therefore \int_S f_n\,d\mu\geq\int_S h_n\,d\mu > K-\frac{1}{n}$

应用MCT，得 $(K\geq)\int_S f\,d\mu\geq K$ [之后证 $\nu-f\mu\perp\nu$ 要用]

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

下面要用Hahn分解定理

<u>Hahn分解定理</u>：设 θ 为 $(S,\mathcal{S})$ 上一个符号测度，则 $\exists G\in\mathcal{S}$ 满足

(i) $\forall A\in\mathcal{S}$ 且 $A\subset G\Rightarrow\theta(A)\geq 0$ (称 G 为 θ 的正集)

(ii) $\forall B\in\mathcal{S}$ 且 $B\subset G^c\Rightarrow\theta(B)\leq 0$ (称 G^c 为 θ 的负集)

以后称 (G,G^c) 为 θ 的Hahn分解

→本质唯一

如果 $(\tilde{G},\tilde{G}^c)$ 也是 θ 的Hahn分解，则 $G\triangle\tilde{G}\,(=G^c\triangle\tilde{G}^c)$ 为 θ-null set.

由于 $\forall n\geq 1$，$\nu_s-\frac{1}{n}\mu$ 为符号测度，故由Hahn定理知 (G_n,G_n^c) 为 $\nu_s-\frac{1}{n}\mu$ 的Hahn分解 $(\forall n\geq 1)$，则

$$
\begin{aligned}
\int_B\left(f+\tfrac{1}{n}1_{G_n}\right)d\mu &= f\mu(B)+\tfrac{1}{n}\mu(B\cap G_n)\\
&= \nu(B)-\nu_s(B)+\tfrac{1}{n}\mu(B\cap G_n)\\
&= \nu(B)-\underbrace{\left(\nu_s(B)-\tfrac{1}{n}\mu(B\cap G_n)\right)}_{\geq\nu_s(B\cap G_n)-\frac{1}{n}\mu(B\cap G_n)\geq 0}\\
&\leq \nu(B)
\end{aligned}
$$

$\Rightarrow f+\frac{1}{n}1_{G_n}\in\mathcal{H}, \forall n\geq 1$

则有如下结论：

(i) $\mu(G_n)=0, \forall n\geq 1$：

(反证)假设 $\exists m\geq 1$, s.t. $\mu(G_m)>0$，故

$\int_S\left(f+\frac{1}{m}1_{G_m}\right)d\mu=\int_S\left(f+\frac{1}{m}\mu(G_m)\right)d\mu>\int_S f\,d\mu\geq K$

但是 $f+\frac{1}{m}1_{G_m}\in\mathcal{H}\Rightarrow\int_S\left(f+\frac{1}{m}1_{G_m}\right)d\mu\leq K$ 于是得到矛盾.

(ii) 定义 $G\triangleq\bigcup_{n=1}^{\infty}G_n$ 故 $\mu(G)=0$. 下证 $\nu_s(G^c)=0$

(反证)假设 $\nu_s(G^c)>0$，故 $\exists n\geq 1$ (充分大)使

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

注意到 $G^c = \bigcap_{n=1}^{\infty} G_n^c \subset G_{n_0}^c$，故 $\nu_s(G^c) - \frac{1}{n}\mu(G^c) \le 0$ } 矛盾！

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

当 μ, ν 为 σ-有限，则 $\exists S$ 的一个划分 $\{B_n\}_{n\ge1}$ 使 $\mu(B_n) < \infty, \nu(B_n) < \infty$

故 $\forall n \ge 1$，定义 $\mu_n \triangleq 1_{B_n}\mu$
$\nu_n \triangleq 1_{B_n}\nu$
$\Rightarrow \mu_n, \nu_n$ 为有限测度且 $\mu = \sum_{n\ge1}\mu_n$，$\nu = \sum_{n\ge1}\nu_n$

注：Hahn 分解定理证明思路：

定义 $\alpha \triangleq \sup_{G\text{为正集}} \theta(G)$。由于 ϕ 是 θ 的正集，故 $\alpha \ge 0$

由"sup"定义，存在 $\{G_n\}_{n\ge1} \subset \mathcal{S}$ 为 θ 的正集，使 $\alpha - \frac{1}{n} < \theta(G_n) \le \alpha \Rightarrow \lim_{n\to\infty}\theta(G_n) = \alpha$

(称 $\{G_n\}_{n\ge1}$ is a maximizing sequence of α)

定义 $G \triangleq \bigcup_{n=1}^{\infty} G_n$，则 $G \in \mathcal{S}$ 以及 G 为 θ 正集，$\theta(G) = \lim_{n\to\infty}\theta(G_n) = \alpha$

下证 G 为 θ 的负集：

(反证) 假设 G^c 不是 θ 的负集，则 $\exists A_0 \subset G^c$ ($A_0 \in \mathcal{S}$) 使 $\theta(A_0) > 0$

$\xrightarrow{\text{Hahn Lemma}} \exists G_0 \subset A_0$ ($G_0 \in \mathcal{S}$)，G_0 为 θ 的正集，使 $\theta(G_0) > 0$

于是定义 $A^* = G_0 \cup G \Rightarrow A^*$ 为 θ 正集

注意到 $G_0 \cap G = \phi$，故 $\theta(A^*) = \theta(G_0) + \theta(G) = \alpha + \theta(G_0) > \alpha$

但 $\theta(A^*) \le \alpha$ 矛盾！

设 θ 为 $(S, \mathcal{S})$ 上一个符号测度，由 Hahn 分解和 (G, G^c) 为 θ 的 Hahn 分解。

定义 $\forall B \in \mathcal{S}$，$\theta^+(B) \triangleq \theta(B \cap G) \ge 0$

$\theta^-(B) \triangleq -\theta(B \cap G^c) \ge 0$

因此 $\underbrace{\theta = \theta^+ - \theta^- \text{ 同时 } \theta^+ \perp \theta^-}_{\theta\text{的 Jordan 分解(唯一)}}$ (事实上，$\theta^+(G^c) = \theta^-(G) = 0$)

称 $|\theta| \triangleq \theta^+ + \theta^-$ 为符号测度的全变差

例. 设μ为$(S,\mathcal{S})$上有限测度以及可测函数$f\in L^1(d\mu)$

定义 $\theta(B)\triangleq f\mu(B)\triangleq\int_B f\,d\mu$

(1) θ的Hahn分解　$G=\{x: f(x)\geq 0\}$，则(G,G^c)为Hahn分解

(2) θ的Jordan分解　$\theta^{\pm}=f^{\pm}\mu$

(3) θ的全变差　$|\theta|=|f|\mu$

定理 (Radon-Nikodym定理):

设μ,ν为$(S,\mathcal{S})$上σ-有限测度，且$\nu\ll\mu$，则$\exists$非负$\mathcal{S}$-可测函数f s.t. $\nu=f\mu$

如果存在非负$\mathcal{S}$-可测函数g s.t. $\nu=g\mu$，则$f=g$, μ-a.e.

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

$$\left.\begin{array}{l}\nu_s\perp\mu\Rightarrow\exists G\in\mathcal{S},\ \text{s.t. } \mu(G)=\nu_s(G^c)=0\\ \nu\ll\mu\Rightarrow\nu(G)=0\Rightarrow\nu_s(G)\leqslant\nu(G)=0\Rightarrow\nu_s(G)=0\end{array}\right\}\Rightarrow\nu_s(S)=0$$

$\Rightarrow\nu_s\equiv 0\Rightarrow\nu=\nu_{ac}=f\mu$

下证条件数学期望的存在唯一性.

定理 (条件数学期望的存在唯一性)

设$X\in L^1(\mathcal{F})$以及$\mathcal{G}\subset\mathcal{F}$为任意$\sigma$-代数，则存在一个可积r.v. $Y\in\mathcal{G}$使

$E[Y1_G]=E[X1_G]$，$\forall G\in\mathcal{G}$

如果存在另一个可积r.v. $\tilde{Y}\in\mathcal{G}$使上述结论成立，则$Y=\tilde{Y}$ a.e.

注: 以后记$E[X|\mathcal{G}]\triangleq Y$. 称$E[X|\mathcal{G}]$为r.v. X关于σ-代数$\mathcal{G}$的条件数学期望

证明: 定义$\mu\triangleq\mathbb{P}|_{\mathcal{G}}$，即$\forall G\in\mathcal{G}$，$\mu(G)=\mathbb{P}(G)$

$\Rightarrow(\Omega,\mathcal{G},\mu)$为一个概率空间 ($\mathcal{G}\subset\mathcal{F}$ σ-代数)

· 假设X是非负可积的，定义$\nu(G)\triangleq E[X1_G]$，$\forall G\in\mathcal{G}$

则ν为$\mathcal{G}$上的一个有限测度 ($\nu(\Omega)=E[X]<\infty$)

$\Rightarrow\nu\ll\mu$ ($\mu(G)=\mathbb{P}(G)=0\Rightarrow\nu(G)=0$)

由R-N定理，$\exists$非负可积$Y\in\mathcal{G}$，s.t. $\nu=Y\mu$ on $\mathcal{G}$

即 $\forall G\in\mathcal{G},\ \nu(G)=\int_G Y\,d\mu=\int_G Y\,d\mathbb{P}=E[Y1_G]$

$\Rightarrow\ E[X1_G]=E[Y1_G]$

• 假设 $X\in L^1(\mathcal{F})$, 令 $X=X^+-X^-$ 即可

• 唯一性: 若 $E[Y1_G]=E[X1_G]=E[\tilde{Y}1_G],\ \forall G\in\mathcal{G}$

$\Rightarrow \forall G\in\mathcal{G},\ E[(Y-\tilde{Y})1_G]=0$

定义 $G_0=\{\omega: Y(\omega)-\tilde{Y}(\omega)>0\}\in\mathcal{G} \Rightarrow Y-\tilde{Y}>0$ on G_0

又 $E[(Y-\tilde{Y})1_{G_0}]=0$ $\Big\}\Rightarrow \mathbb{P}(G_0)=0$

$\tilde{G}_0\triangleq\{\tilde{Y}-Y>0\}\in\mathcal{G}$ $\quad \tilde{Y}-Y>0$ on $\tilde{G}_0$

$E[(\tilde{Y}-Y)1_{\tilde{G}_0}]=0$ $\Big\}\Rightarrow \mathbb{P}(\tilde{G}_0)=0$

$\Rightarrow \mathbb{P}(X=Y)=1-\mathbb{P}(G_0)-\mathbb{P}(\tilde{G}_0)=1$

§2. 条件数学期望性质

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

(C1) $X\in L^1(\mathcal{G})$ (X 可积且 $X\in\mathcal{G}$), 则 $E[X|\mathcal{G}]=X$

特别地, $X\equiv C$ (常数), 有 $E[C|\mathcal{G}]=C$

$Z\perp\mathcal{G}\Rightarrow E[Z|\mathcal{G}]=E[Z]$ (独立)

(C2) 设 $Y\in\mathcal{G}$ ($Y:\Omega\to S_Y$), $Z\in\mathcal{F}$ 且 $Z\perp\mathcal{G}$ (即 $\sigma(Z)$ 与 $\mathcal{G}$ 独立), $Z:\Omega\to S_Z$

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

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

证明: (i) $g(Y)\in\mathcal{G}$

(ii) $\forall G\in\mathcal{G},\ E[\varphi(Y,Z)1_G]=E[g(Y)1_G]$

事实上, 令 $X\triangleq 1_G$, 故 $X\in\mathcal{G}$

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

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

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

$$= \int x \left(\int \varphi(y,z) P_Z(dz)\right) P_{(X,Y)}(dx,dy)$$

$$\underset{\text{formula}}{\overset{\text{change of variable}}{=\!=\!=}} \int x\, g(y)\, P_{(X,Y)}(dx,dy)$$

$$\underset{\text{formula}}{\overset{\text{change of variable}}{=\!=\!=}} E[g(Y)X]$$

注：$\varphi(y,z)=f(y)h(z)$，其中 f,h 均可测. 故 $E[f(Y)h(Z)|\mathcal{G}]=g(Y)$

其中 $g(y)=E[\varphi(y,Z)]=f(y)E[h(Z)]$

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

若 $f\equiv 1 \Rightarrow E[h(Z)|\mathcal{G}]=E[h(Z)]$

若 $h\equiv 1 \Rightarrow E[f(Y)|\mathcal{G}]=f(Y)$

(C3) $X\in L^1(\mathcal{F})$，则 $E\{E[X|\mathcal{G}]\}=E[X]$

事实上，由 $E\{E[X|\mathcal{G}]1_G\}=E[X1_G]$，$\forall G\in\mathcal{G}$

取 $G=\Omega$，得 $E\{E[X|\mathcal{G}]\}=E[X]$

例. $X \triangleq \sum_{i=1}^{N}\xi_i$，$\xi_1,\cdots,\xi_n,\cdots$ i.i.d，且 $E[\xi_i]=\mu\in\mathbb{R}$，$Var[\xi_i]=\sigma^2>0$

这里 $N\sim Poisson(\lambda)$，$\lambda>0$ 且与 ξ_n 相互独立. 计算 $E[X]$，$Var[X]$

解：$E\{E[X|\mathcal{G}]\}=E[X]$

$$\Downarrow \mathcal{G}=\sigma(N)$$

$$E[X]=E\{E[X|\sigma(N)]\}=E\{\underbrace{E[\textstyle\sum_{i=1}^{N}\xi_i|\sigma(N)]}_{\in\sigma(N)}\}$$

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

进一步 $f(n)=E[\sum_{i=1}^{n}\xi_i]=nE[\xi_1]=n\mu$

故 $E[X]=E[f(N)]=E[\mu N]=\mu E[N]=\lambda\mu$

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

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

(C4) $X\in L^1(\mathcal{F})$，如果 $X\geq 0$ a.e.，则 $E[X|\mathcal{G}]\geq 0$ a.e.

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

证：由等式，$\forall G \in \mathcal{G}$，$E[X1_G] = E\{E[X|\mathcal{G}]1_G\}$

假设 $X \geq 0$ a.e.，则 $E[X1_G] \geq 0$，$\forall G \in \mathcal{G}$

因此 $E\{E[X|\mathcal{G}]1_G\} \geq 0$，$\forall G \in \mathcal{G}$　$(*)$

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

由于 $E[X|\mathcal{G}] \in \mathcal{G} \Rightarrow G_0 \in \mathcal{G} \xRightarrow{(*)} E\{E[X|\mathcal{G}]1_{G_0}\} \geq 0$

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

于是 $E\{E[X|\mathcal{G}]1_{G_0}\} = 0 \Rightarrow \left.\begin{array}{l} E\{(-E[X|\mathcal{G}])1_{G_0}\} = 0 \\ -E[X|\mathcal{G}] > 0 \text{ on } G_0 \end{array}\right\} \Rightarrow \mathbb{P}(G_0) = 0$

(C5) $X \in L^1(\mathcal{F})$，$Y \in L^1(\mathcal{F})$，$\alpha, \beta \in \mathbb{R}$，则 $E[\alpha X + \beta Y|\mathcal{G}] = \alpha E[X|\mathcal{G}] + \beta E[Y|\mathcal{G}]$

证：$E\{E[X|\mathcal{G}]1_G\} = E[X1_G]$

$E\{E[Y|\mathcal{G}]1_G\} = E[Y1_G]$

因此 $E\{(\alpha E[X|\mathcal{G}] + \beta E[Y|\mathcal{G}])1_G\} = E[(\alpha X + \beta Y)1_G]$，$\forall G \in \mathcal{G}$

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

(C6) 设 $X \in L^1(\mathcal{F})$ 以及 $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$（$\sigma$-代数）

$E\{E[X|\mathcal{G}]|\mathcal{H}\} = E[X|\mathcal{H}]$　(Tower property)

证：设 $Y \triangleq E[X|\mathcal{G}]$，$Z \triangleq E[Y|\mathcal{H}]$

由 $Y = E[X|\mathcal{G}] \Rightarrow E[X1_G] = E[Y1_G]$，$\forall G \in \mathcal{G}$

由 $Z = E[Y|\mathcal{H}] \Rightarrow E[Y1_H] = E[Z1_H]$，$\forall H \in \mathcal{H}$

由于 $\mathcal{H} \subset \mathcal{G}$，故 $\forall H \in \mathcal{H} \Rightarrow H \in \mathcal{G} \Rightarrow E[Z1_H] = E[X1_H]$，$\forall H \in \mathcal{H}$

$\Rightarrow Z = E[X|\mathcal{H}]$

注：条件数学期望的不等式及收敛定理，

- $E[|XY|] \leq \{E[|X|^p]\}^{\frac{1}{p}} \cdot \{E[|Y|^q]\}^{\frac{1}{q}}$，$1 < p < +\infty$，$\frac{1}{p} + \frac{1}{q} = 1$，$X \in L^p(\mathcal{F})$，$Y \in L^q(\mathcal{F})$

特别地，$p = q = 2$ 时，$E|XY| \leq \sqrt{E|X|^2 \cdot E|Y|^2}$

令 $H \triangleq L^2(\mathcal{F}) \triangleq \{\text{r.v. } X \in \mathcal{F} : E|X|^2 < +\infty\}$，$\|X\|_{L^2} \triangleq (E|X|^2)^{\frac{1}{2}}$

$\Rightarrow (H, \|\cdot\|_{L^2})$ 为 Hilbert 空间

定义内积 $\langle X, Y\rangle = E[XY]$, $X, Y \in H$, 则有 $|\langle X, Y\rangle| \le \|X\|_{L^2} \cdot \|Y\|_{L^2}$

$$E[|XY| \mid \mathcal{G}] \le \{E[|X|^p \mid \mathcal{G}]\}^{\frac{1}{p}} \cdot \{E[|Y|^q \mid \mathcal{G}]\}^{\frac{1}{q}}$$

- Jensen 不等式: 设 $X \in \mathcal{F}$ 以及 $\varphi: \mathbb{R} \to \mathbb{R}$ 凸函数, 且使 $\varphi(X)$ 可积, 则

$$\varphi(E[X]) \le E[\varphi(X)]$$
$$\varphi(E[X \mid \mathcal{G}]) \le E[\varphi(X) \mid \mathcal{G}]$$

例. $\varphi(x) = x^2$, 由 Jensen 不等式 $(EX)^2 \le EX^2$

- 单调收敛定理: 设 $\{X_n\}_{n\ge1}$ 为一列递增非负(可积) r.v.s, 则

$$\lim_{n\to\infty} E[X_n \mid \mathcal{G}] = E[\lim_{n\to\infty} X_n \mid \mathcal{G}]$$

Fatou 引理: 设 $\{X_n\}_{n\ge1}$ 为一列非负(可积) r.v.s, 则

$$E[\varliminf_{n\to\infty} X_n \mid \mathcal{G}] \le \varliminf_{n\to\infty} E[X_n \mid \mathcal{G}]$$

控制收敛定理: 设 $\{X_n\}_{n\ge1}$ 为一列 r.v. 且 $|X_n| \le Y$ a.e., 其中 $Y \in L^1(\mathcal{F})$, 则

$$\lim_{n\to\infty} E[X_n \mid \mathcal{G}] = E[\lim_{n\to\infty} X_n \mid \mathcal{G}]$$

练习: 设 $X \in L^1(\mathcal{F})$, $\mathcal{G} \subset \mathcal{F}$ (σ-代数), 则 $X = E[X \mid \mathcal{G}] + \xi$

其中 r.v. ξ 满足 (1) $E[\xi] = 0$ (2) $E[\xi Y] = 0$, $\forall Y \in L^2(\mathcal{G})$

(1) $E\xi = E[X - E[X \mid \mathcal{G}]] = EX - E[E[X \mid \mathcal{G}]] = EX - EX = 0$

(2) $\forall G \in \mathcal{G}$, $E[X 1_G] = E\{E[X \mid \mathcal{G}] 1_G\} \Rightarrow E[\xi 1_G] = 0$, $\forall G \in \mathcal{G}$

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

- 若 $Y \in L^2(\mathcal{G})$ 且 Y 为简单 r.v. (即 $\exists \{G_n\}_{n\ge1} \subset \mathcal{G}$ 为 Ω 的一个划分

  s.t. $Y = \sum_{n=1}^{m} a_n 1_{G_n}$, $m = 1, 2, \cdots, +\infty$, $a_n > 0$)

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

  $m < +\infty$, $E[\xi^+ \sum_{n=1}^{m} a_n 1_{G_n}] = E[\xi^- \sum_{n=1}^{m} a_n 1_{G_n}]$

  $m = +\infty$, MCT $\Rightarrow E[\xi^+ \sum_{n=1}^{\infty} a_n 1_{G_n}] = E[\xi^- \sum_{n=1}^{\infty} a_n 1_{G_n}]$

  此即对任意 $Y \in L^2(\mathcal{G})$ 且 Y 非负简单函数有 $E[\xi^+ Y] = E[\xi^- Y]$ ($\Leftrightarrow E[\xi Y] = 0$)

- $Y \in L^2(\mathcal{G})$ 且 Y 非负, 则 $\exists$ 一列非负简单 $L^2(\mathcal{G})$-r.v.s $\{Y_n\}_{n\ge1}$ s.t.

$Y_n(\omega)\uparrow Y(\omega),\ \forall \omega\in\Omega$

由上一步有 $E[\xi^+ Y_n]=E[\xi^- Y_n],\ \forall n\geq 1$

则由MCT $\Rightarrow E[\xi^+ Y]=E[\xi^- Y] \Leftrightarrow E[\xi Y]=0$

· $Y\in L^2(\mathcal{G})$, 考虑 $Y=Y^+-Y^-$, 由上一步有 $\left.\begin{array}{l} E[\xi Y^+]=0 \\ E[\xi Y^-]=0 \end{array}\right\} \Rightarrow E[\xi Y]=0$

§3. 条件数学期望的应用

<u>习题</u>: 设 $Y\in L^2(\mathcal{F})$ 以及 $\mathcal{G}\subset\mathcal{F}$ (σ-代数), 考虑 $V_Y \triangleq \inf\limits_{\eta\in L^2(\mathcal{G})} E[|Y-\eta|^2]=E[|Y-\eta^*|^2]$

问: $\eta^*=?$

分析: $Y\in L^2(\mathcal{G}) \xrightarrow[\text{or Hölder不等式}]{\text{Jensen不等式}} E[Y|\mathcal{G}]\in L^2(\mathcal{G})$

事实上, $\forall \eta\in L^2(\mathcal{G})$

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

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

下证交叉项为0:

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

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

$$=E\{Z\underbrace{E[Y-E[Y|\mathcal{G}]|\mathcal{G}]}_{=0}\}=0$$

因此 $\forall \eta\in L^2(\mathcal{G})$, 有

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

$$\geq E[|Y-E[Y|\mathcal{G}]|^2]$$

等号成立当且仅当 $\eta=\eta^*\triangleq E[Y|\mathcal{G}]$

例. 统计决策问题

输入变量 $X\in\mathbb{R}^n$ (r.v.)

输出变量 $Y\in\mathbb{R}$ (r.v.)

能否找到一个可测函数 $f:\mathbb{R}^n\to\mathbb{R}$, 使 $f(x)$ 在某种意义下可以逼近 Y, 即 $Y\approx f(x)$

$\longrightarrow$ 数学问题: 设 $L:\mathbb{R}\times\mathbb{R}\to\mathbb{R}_+$ 为一可测函数 (损失函数或成本函数 loss fcn or cost fcn). 考虑 $\inf\limits_{f\in\mathcal{B}(\mathbb{R}^n\to\mathbb{R})} E[L(Y,f(X))]$

特别地,取 $L(y_1,y_2)=|y_1-y_2|^2$, $y_1,y_2\in\mathbb{R}$

又注意到 $f\in\mathcal{B}(\mathbb{R}^n;\mathbb{R})\Rightarrow f(X)\in\sigma(X)$ (取 $\mathcal{G}\triangleq\sigma(X)$)

考虑 $\inf\limits_{\eta\in L^2(\sigma(X))} E[|Y-\eta|^2]=E[|Y-\eta^*|^2]$,

其中 $\eta^*=E[Y|\sigma(X)]\in\sigma(X)\Rightarrow\exists f^*\in\mathcal{B}(\mathbb{R}^n\to\mathbb{R})$ s.t. $\eta^*=f^*(X)$, 其中 $f^*(x)=E[Y|X=x]$ $x\in\mathbb{R}^n$

称为回归函数