3.条件期望

1.基本定义

定义:

$$\text{设}(\Omega,\mathcal{F},\mathbb{P})\text{是一概率空间，}\mathcal{G}\text{是}\mathcal{F}\text{中}\sigma\text{代数，}X\text{是}\Omega\text{上定义的（实）随机变量，}\mathbb{E}[X]\text{存在，在}\mathcal{G}\text{上定义}$$
$$\nu(A)=\int_A Xd\mathbb{P}\text{，}A\in\mathcal{G}$$
$$\text{则}\nu\text{是}\mathcal{G}\text{上符号测度，且}\nu\ll\mathbb{P}_{\mathcal{G}}(\mathbb{P}_{\mathcal{G}}\text{表示}\mathbb{P}\text{限制在}\mathcal{G}\text{上}),$$
$$\text{故}\frac{d\nu}{d\mathbb{P}_{\mathcal{G}}}\text{存在，称之为给定}\mathcal{G}\text{之下}X\text{的条件期望，记为}\mathbb{E}[X|\mathcal{G}]$$

关于该定义有以下几个注解:

$$\text{注}1:\nu(A)=\int_A\mathbb{E}[X|\mathcal{G}]d\mathbb{P}_{\mathcal{G}}$$
$$\text{注}2:\mathbb{E}[X|\mathcal{G}]\text{是}\mathcal{G}\text{可测随机变量}$$
$$\text{注}3\text{：}\mathbb{E}[X|\mathcal{G}]\text{是}\mathbb{P}_{\mathcal{G}}\ \ a.s\text{唯一（在概率空间}(\Omega,\mathcal{G},\mathbb{P}_{\mathcal{G}})\text{上）}$$
$$\text{即}\nu(A)=\int_A y_1d\mathbb{P}_{\mathcal{G}}=\int_A y_2d\mathbb{P}_{\mathcal{G}},\forall A\in\mathcal{G}\Rightarrow y_1=y_2(\mathbb{P}_{\mathcal{G}}\ a.s)$$
$$\text{注}4:\int_A\mathbb{E}[X|\mathcal{G}]d\mathbb{P}=\int_A Xd\mathbb{P},\forall A\in\mathcal{G}$$

如果$X=1_A,A\in\mathcal{F}$，那么

$$\mathbb{E}[X|\mathcal{G}]=\mathbb{P}(A|\mathcal{G})$$

若Y是$(\Omega,\mathcal{F},\mathbb{P})$上定义，取值于可测空间$(E,\Sigma)$中的随机变量（随机元），则

$$\mathbb{E}[X|\sigma(Y)]=\mathbb{E}[X|Y]$$
$$\text{称为给定}Y\text{之下}X\text{的条件期望}$$
$$\text{由}\sigma(Y)\text{可测，随机变量的构造知，}$$
$$\text{存在可测映射}\varphi:(E,\Sigma)\to(\mathbb{R},\mathcal{B})\text{使得}\mathbb{E}[X|Y]=\mathbb{E}[X|\sigma(Y)]=\varphi(Y)$$
$$\text{记}\mathbb{E}[X|Y=y]=\varphi(y)\text{，称之为给定}Y=y\text{条件下}X\text{的条件期望}$$

引理

$$f,g\text{都是}(\Omega,\mathcal{F},\mathbb{P})\text{上可测函数且}\int_A fd\mathbb{P}=\int_A gd\mathbb{P},\forall A\in\mathcal{F}$$
$$\text{则}f=g(a.s)$$

证明：利用反证法。记

$$A=\{f\neq g\}=\{f>g\}\bigcup\{f<g\}$$

现在假定$\mathbb{P}(A)>0$，那么$\mathbb{P}(f>g)>0$和$\mathbb{P}(f<g)>0$其中之一必成立，不妨假设$\mathbb{P}(f>g)>0$，注意到

$$\{f>g\}=\bigcup_{n=1}^{\infty}\{f>g+\frac{1}{n}\}$$

所以存在n_0使得

$$\mathbb{P}(f>g+\frac{1}{n_0})>0$$

从而

$$\begin{aligned}\int_{f>g+\frac{1}{n_0}}fd\mathbb{P}&>\int_{f>g+\frac{1}{n_0}}(g+\frac{1}{n_0})\mathbb{P}\\&=\int_{f>g+\frac{1}{n_0}}gd\mathbb{P}+\frac{1}{n_0}\int_{f>g+\frac{1}{n_0}}d\mathbb{P}\\&>\int_{f>g+\frac{1}{n_0}}gd\mathbb{P}\end{aligned}$$

而$\{f>g+\frac{1}{n_0}\}$可测，这就与假定矛盾。

命题

$$\begin{gathered}\text{设 }X\text{是 }(\Omega,\mathcal{F},\mathbb{P})\text{上定义的（实）}r.v,\mathbb{E}[X]\text{存在，}\\\mathcal{G}\text{是 }\mathcal{F}\text{上的}\sigma\text{代数，}Y\text{是 }\Omega\text{上定义的实值映射，若它满足：}\\(1)Y\text{是 }\mathcal{G}\text{可测}\\(2)\forall A\in\mathcal{G},\int_A Yd\mathbb{P}=\int_A Xd\mathbb{P}\\\text{则 }Y=\mathbb{E}[X|\mathcal{G}](a.s)\end{gathered}$$

证明：由条件知

$$\forall A\in\mathcal{G},\int_A Yd\mathbb{P}=\int_A\mathbb{E}[X|\mathcal{G}]d\mathbb{P}=\int_A Xd\mathbb{P}$$

由引理可得

$$Y=\mathbb{E}[X|\mathcal{G}](a.s)$$

2.基本性质

$$\begin{aligned}&(1)\text{若}\mathbb{E}[X]\text{存在，则}\forall\text{常数 }a,\ \mathbb{E}[aX|\mathcal{G}]=a\mathbb{E}[X|\mathcal{G}](a.s)\\&(2)\text{若}\mathbb{E}[X],\ \mathbb{E}[Y]\text{存在，且}\mathbb{E}[X]+\mathbb{E}[Y]\text{有意义，则}\mathbb{E}[X+Y|\mathcal{G}]=\mathbb{E}[X|\mathcal{G}]+\mathbb{E}[Y|\mathcal{G}](a.s)\\&(3)\text{若}\mathbb{E}[X],\ \mathbb{E}[Y]\text{存在且}X\le Y(a.s),\text{则}\mathbb{E}[X|\mathcal{G}]\le\mathbb{E}[Y|\mathcal{G}](a.s)\\&(4)\text{若}0\le X_n\uparrow X,\text{则}0\le\mathbb{E}[X_n|y]\uparrow\mathbb{E}[X|y](a.s)\\&(5)\text{若}X_n\ge Y,Y\text{可积，则}\mathbb{E}[\varliminf_{n\to\infty}X_n|Y]\le\varliminf_{n\to\infty}\mathbb{E}[X_n|Y](a.s)\\&(6)\text{若}\forall n\ge1,|X_n|\le Y,Y\text{可积且}X_n\to X,\text{则}\lim_{n\to\infty}\mathbb{E}[X_n|\mathcal{G}]=\mathbb{E}[X|\mathcal{G}]=\mathbb{E}[\lim_{n\to\infty}X_n|\mathcal{G}](a.s)\\&(7)\text{若}X\text{可积，则}\mathbb{E}[X|\mathcal{G}]\text{可积且}\mathbb{E}[\mathbb{E}[X|\mathcal{G}]]=\mathbb{E}[X](a.s)\end{aligned}$$

证明：(1)$a\mathbb{E}[X|\mathcal{G}]$是$\mathcal{G}$可测，且$\forall A\in\mathcal{G}$，我们有

$$
\begin{aligned}
\int_A a\mathbb{E}[X|\mathcal{G}]d\mathbb{P} &= a\int_A \mathbb{E}[X|\mathcal{G}]d\mathbb{P} \\
&= a\int_A Xd\mathbb{P} \\
&= \int_A aXd\mathbb{P}
\end{aligned}
$$

(2)因为$\mathbb{E}[X|\mathcal{G}]$，$\mathbb{E}[Y|\mathcal{G}]$都可测，所以$\mathbb{E}[X|\mathcal{G}]+\mathbb{E}[Y|\mathcal{G}]$可测，注意到

$$
\begin{aligned}
\int_A （\mathbb{E}[X|\mathcal{G}]+\mathbb{E}[Y|\mathcal{G}]） d\mathbb{P} &= \int_A \mathbb{E}[X|\mathcal{G}]d\mathbb{P} + \int_A \mathbb{E}[Y|\mathcal{G}]d\mathbb{P} \\
&= \int_A Xd\mathbb{P} + \int_A Yd\mathbb{P} \\
&= \int_A (X+Y)d\mathbb{P}
\end{aligned}
$$

(3)$\forall A \in \mathcal{G}$，我们有

$$
\int_A \mathbb{E}[X|\mathcal{G}]d\mathbb{P} = \int_A Xd\mathbb{P} \le \int_A Yd\mathbb{P} = \int_A \mathbb{E}[Y|\mathcal{G}]d\mathbb{P}
$$

所以

$$
\mathbb{E}[X|\mathcal{G}] \le \mathbb{E}[Y|\mathcal{G}]
$$

(4)由(3)可得$\mathbb{E}[X_n|\mathcal{G}]\uparrow$，记

$$
Y = \lim_{n\to\infty} \mathbb{E}[X_n|\mathcal{G}]
$$

注意到$\forall A \in \mathcal{G}$

$$
\begin{aligned}
\int_A Yd\mathbb{P} &= \int_A \lim_{n\to\infty} \mathbb{E}[X_n|\mathcal{G}]d\mathbb{P} \\
&\overset{\text{Levi}}{=} \lim_{n\to\infty} \int_A \mathbb{E}[X_n|\mathcal{G}]d\mathbb{P} \\
&= \lim_{n\to\infty} \int_A X_n d\mathbb{P} \\
&= \int_A \lim_{n\to\infty} X_n d\mathbb{P} \\
&= \int_A Xd\mathbb{P}
\end{aligned}
$$

所以

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

(5)记

$$
0 \le X_n - Y \triangleq \hat{X}_n
$$

注意到

$$
\varliminf_{n\to\infty} \hat{X}_n = \varliminf_{n\to\infty} (X_n - Y) = \lim_{n\to\infty} \inf_{m\ge n} (X_m - Y)
$$

所以

$$
\begin{aligned}
\mathbb{E}[\varliminf_{n\to\infty}(X_n-Y)|\mathcal{G}] &= \mathbb{E}[\lim_{n\to\infty}\inf_{m\geq n}(X_m-Y)|\mathcal{G}] \\
&= \lim_{n\to\infty}\mathbb{E}[\inf_{m\geq n}(X_m-Y)|\mathcal{G}] \\
&\leq \lim_{n\to\infty}\inf_{m\geq n}\mathbb{E}[(X_m-Y)|\mathcal{G}] \\
&= \varliminf_{n\to\infty}\mathbb{E}[(X_n-Y)|\mathcal{G}](a.s)
\end{aligned}
$$

从而

$$
\mathbb{E}[\varliminf_{n\to\infty}X_n|\mathcal{G}]-\mathbb{E}[Y|\mathcal{G}]\leq\varliminf_{n\to\infty}\mathbb{E}[X_n|\mathcal{G}]-\mathbb{E}[Y|\mathcal{G}]
$$

因为Y可积，从而消去$\mathbb{E}[Y|\mathcal{G}]$可得

$$
\mathbb{E}[\varliminf_{n\to\infty}X_n|\mathcal{G}]\leq\varliminf_{n\to\infty}\mathbb{E}[X_n|\mathcal{G}]
$$

(6)对X_n以及$-X_n$应用(5)即可证明(6)

(7)证略，可以得到如下结果

$$
\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}[\mathbb{E}[X|y]] \\
&= \mathbb{E}[\varphi(y)] \\
&= \int_{\Omega}\varphi(y)d\mathbb{P} \\
&= \int_{\mathbb{R}}\varphi(y)dF_y \\
&= \int_{\mathbb{R}}\mathbb{E}[\mathbb{E}[X|Y=y]]\mathbb{P}(y\in dy)
\end{aligned}
$$

例1

设 (X,Y)是连续型随机变量，若有联合密度$p(x,y)$，边缘密度$p_1(x),p_2(y)$，记

$$
p_{X|Y}(x,y)=\begin{cases}\frac{p(x,y)}{p_2(y)} & p_2(y)>0\\ 0 & p_2(y)=0\end{cases}
$$

再记

$$
\varphi(y)=\mathbb{E}[X|Y=y]=\int_{-\infty}^{+\infty}xp_{X|Y}(x,y)dx
$$

则$\varphi(Y)=\mathbb{E}[X|Y](a.s)$

事实上，$\forall A=Y^{-1}(B),B\in\mathcal{B}$，即$A=\{w:Y(w)\in\mathcal{B}\}$

$$
\begin{aligned}
\int_A \varphi(y)d\mathbb{P} &= \int_{\{Y\in B\}} \varphi(y)d\mathbb{P} \\
&= \int_\Omega \varphi(y)1_{\{Y\in B\}}(w)d\mathbb{P} \\
&= \int_\Omega \varphi(y)1_B(y)d\mathbb{P} \\
&= \mathbb{E}[\varphi(y)1_B(y)] \\
&= \int_{-\infty}^{+\infty} \varphi(y)1_B(y)p_2(y)dy \\
&= \int_{p_2(y)=0} \varphi(y)1_B(y)p_2(y)dy + \int_{p_2(y)>0} \varphi(y)1_B(y)p_2(y)dy \\
&= \int_{p_2(y)>0} dy\int_{-\infty}^{+\infty} xp_{X|Y}(x,y)p_2(y)1_B(y)dx \\
&= \int_{-\infty}^{+\infty} dy\int_{-\infty}^{+\infty} xp(x,y)1_B(y)dx \\
&= \int_\Omega x1_B(y)d\mathbb{P} \\
&= \int_{\{Y\in B\}} xd\mathbb{P} \\
&= \int_A xd\mathbb{P}
\end{aligned}
$$

命题

若B是$\mathcal{G}$的非空原子（即$C\subset B\Rightarrow C=\varnothing$或$C=B$），
则在B上，$\mathbb{E}[X|\mathcal{G}]$是常数，当$\mathbb{P}(B)>0$时，这个常数为

$$
\mathbb{E}[X|B]\triangleq\frac{\mathbb{E}[X1_B]}{\mathbb{P}(B)}
$$

证明：由于B非空，则$\exists w_0\in B$，记

$$
B^*=\{w\in B:\mathbb{E}[X|\mathcal{G}](w)=\mathbb{E}[X|\mathcal{G}](w_0)\}
$$

则$B^*\in\mathcal{G}$且$B^*\subset B$，因为B是原子，所以

$$
B^*=B
$$

注意到

$$
\mathbb{P}(B)\mathbb{E}[X|B]=\int_B\mathbb{E}[X|\mathcal{G}]d\mathbb{P}=\int_B Xd\mathbb{P}
$$

所以该常数为

$$
\mathbb{E}[X|B]=\frac{\mathbb{E}[X1_B]}{\mathbb{P}(B)}
$$

习题

习题1

（课本P89/7.3/1）

解：

$$
\begin{aligned}
\mathbb{P}(Y=m) &= \sum_{m=n}^{\infty} \mathbb{P}(Y=m|X=n)\mathbb{P}(X=n) \\
&= \sum_{m=n}^{\infty} C_n^m p^m (1-p)^{n-m} \frac{e^{-\lambda}\lambda^n}{n!} \\
&= \sum_{m=n}^{\infty} \frac{n!}{m!(n-m)!} p^m (1-p)^{n-m} \frac{e^{-\lambda}\lambda^n}{n!} \\
&= \frac{e^{-\lambda}(\lambda p)^m}{m!} \sum_{n=m}^{\infty} \frac{(\lambda-\lambda p)^{n-m}}{(n-m)!} \\
&= \frac{e^{-\lambda}(\lambda p)^m}{m!} e^{\lambda-\lambda p} \\
&= \frac{e^{-\lambda p}(\lambda p)^m}{m!}
\end{aligned}
$$

习题2

（课本P89/7.3/2）

解：总时间$T = NX_k$，所以

$$
\begin{aligned}
\mathbb{E}[T] &= \mathbb{E}[NX_k] \\
&= \mathbb{E}[\mathbb{E}[NX_k|N=n]] \\
&= \sum_{n=0}^{\infty} \frac{e^{-\lambda}(\lambda)^n}{n!} n\mathbb{E}[X_k] \\
&= \frac{\lambda}{\beta} \sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^{n-1}}{(n-1)!} \\
&= \frac{\lambda}{\beta}
\end{aligned}
$$