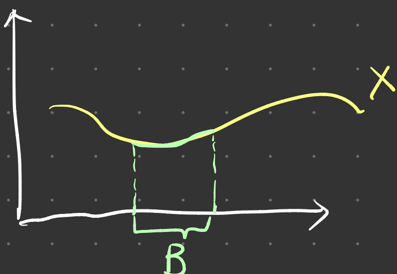
2.1 Conditional probability and conditional expectation

一、条件概率与条件期望的定义：

$$P(A|B)=\frac{P(A\cap B)}{P(B)}=\frac{P(AB)}{P(B)}$$

$E(\cdot|B)$ 可以看作到 B 集合上 X 的平均高度（其实是面积，但底为1，所以可以直接看作高度



例：$\Omega=\{1,2,3,4\}$　$\mathcal{F}=\sigma(\{1\},\{2\},\{3\},\{4\})$

$P(\{1\})=\frac{1}{2}$，$P(\{2\})=\frac{1}{4}$，$P(\{3\})=\frac{1}{6}$，$P(\{4\})=\frac{1}{12}$

$B=\{1,2,3\}$

$$P(A|B)=\frac{P(AB)}{P(B)}=\frac{P(AB)}{\frac{1}{2}+\frac{1}{4}+\frac{1}{6}}=\frac{P(AB)}{\frac{6+3+2}{12}}=\frac{12}{11}P(AB)$$

$$\Rightarrow P(\{1\}|B)=\frac{12}{11}P(\{1\})=\frac{12}{11}\times\frac{1}{2}=\frac{6}{11}$$

Let $X(1)=4$, $X(2)=3$, $X(3)=2$, $X(4)=1$

求 $E(X)$

$$E(X)=4\times\frac{1}{2}+3\times\frac{1}{4}+2\times\frac{1}{6}+1\times\frac{1}{12}=2+\frac{3}{4}+\frac{1}{3}+\frac{1}{12}=\frac{24+9+4+1}{12}=\frac{38}{12}=\frac{19}{6}$$

求 $E(X|B)$

$$E(X|B)=4\times P(\{1\}|B)+3\times P(\{2\}|B)+2\times P(\{3\}|B)+1\times P(\{4\}|B)$$

二、随机变量的条件期望

离散情况：

Let X and Y be discrete r.v. on $(\Omega,\mathcal{F},\mathbb{P})$ and assume that $P(Y=y)>0$.

(1) The function:

$$P_{X|Y}(x|y)=\mathbb{P}(X=x\mid Y=y)=\frac{P(X=x,Y=y)}{P(Y=y)}=\frac{p(x,y)}{P_Y(y)}$$

$$P_Y(y)=\sum_x p(x,y)$$

is called the conditional probability mass function of X given $Y=y$

(2) The conditional expectation of X given that $Y=y$

$$E(X\mid Y=y)=\sum_x x\,P_{X|Y}(x,Y=y)=\sum_x\frac{x\,p(x,y)}{P_Y(y)}$$

条件期望实际上就是把整个 sample space 缩小.

例子: $\Omega=\{1,2,3\}$

$\mathcal{F}=\sigma(\{1\},\{2\},\{3\})$

$\mathbb{P}(\{1\})=\frac{1}{2}$, $\mathbb{P}(\{2\})=\frac{1}{3}$, $\mathbb{P}(\{3\})=\frac{1}{6}$

(1) Suppose $X(\omega)=\omega$, $Y(\omega)=4-\omega$.

Then $p(1,1)=\mathbb{P}(\{\omega: X(\omega)=1, Y(\omega)=1\})=\mathbb{P}(\emptyset)=0$

$p(1,2)=\mathbb{P}(\{\omega: X(\omega)=1, Y(\omega)=2\})=\mathbb{P}(\emptyset)=0$

$p(1,3)=\mathbb{P}(\{\omega: X(\omega)=1, Y(\omega)=3\})=\mathbb{P}(\{1\})=\frac{1}{2}$

$p(2,1)=p(2,3)=0$

$p(2,2)=\frac{1}{3}$

$p(3,1)=\frac{1}{6}$

$p(3,2)=p(3,3)=0$

$p_Y(Y=1)=\mathbb{P}(\{\omega: Y(\omega)=1\})=\mathbb{P}(\{3\})=\frac{1}{6}$

$p_Y(Y=2)=\mathbb{P}(\{\omega: Y(\omega)=2\})=\mathbb{P}(\{2\})=\frac{1}{3}$

$p_Y(Y=3)=\mathbb{P}(\{\omega: Y(\omega)=3\})=\mathbb{P}(\{1\})=\frac{1}{2}$

$p_{X|Y}(1|1)=\frac{p(1,1)}{p_Y(1)}=0$, $p_{X|Y}(1|2)=0$, $p_{X|Y}(1|3)=\frac{p(1,3)}{p_Y(3)}=\frac{\frac{1}{2}}{\frac{1}{2}}=1$

(2) Suppose $p(1,1)=\frac{1}{36}$, $p(1,2)=\frac{1}{4}$, $p(1,3)=\frac{1}{48}$

$p(2,1)=\frac{1}{12}$, $p(2,2)=\frac{1}{24}$, $p(2,3)=\frac{1}{3}$

$p(3,1)=\frac{1}{48}$, $p(3,2)=\frac{1}{3}$, $p(3,3)=\frac{1}{6}$

可求 $P_Y(1)$, $P_Y(2)$, $P_Y(3)$

连续情况：

连续时无法考虑 $Y=y$ 某一点的概率情况，我们转而考虑区间（用 cdf）

$$P(x\le X\le x+\Delta x \mid y\le Y\le y+\Delta y)$$

$$=\frac{P(x\le X\le x+\Delta x,\ y\le Y\le y+\Delta y)}{P(y\le Y\le y+\Delta y)}$$

$$=\frac{\int_y^{y+\Delta y}\int_x^{x+\Delta x} f(\tilde{x},\tilde{y})\,d\tilde{x}\,d\tilde{y}}{\int_y^{\Delta y} f_Y(\tilde{y})\,d\tilde{y}} \xrightarrow[\Delta y\to 0]{} \frac{\int_x^{x+\Delta x} f(\tilde{x},y)\,d\tilde{x}}{f_Y(y)} \xrightarrow[\Delta x\to 0]{} \frac{f(x,y)}{f_Y(y)}$$

$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}$$

$$E[X|Y=y]=\int_{-\infty}^{+\infty} x\, f_{X|Y}(x|y)\,dx \quad (f_Y(y)>0)$$

例：Suppose the joint density function of X and Y is given by

$$f(x,y)=\frac{1}{y}e^{-\frac{x}{y}}e^{-y}\quad,\ 0<x,y<\infty$$

Compute $E[X|Y=y]$ and $E[X^2|Y=y]$

解：$E[X|Y=y]=\int_0^{+\infty}x\,f_{X|Y}(x|Y=y)\,dx=\int_0^{+\infty}x\cdot\frac{f(X=x,Y=y)}{f_Y(y)}dx$

$$=\int_0^{+\infty}x\cdot\frac{f(x,y)}{\int_0^{+\infty}f(x,y)\,dx}dx=\int_0^{+\infty}x\cdot\frac{f(x,y)}{e^{-y}}dx$$

$$=\int_0^{+\infty}\frac{x}{y}e^{-\frac{x}{y}}dx=y$$

$$E[X^2|Y=y]=\int_0^{+\infty}x^2\cdot\frac{f(x,y)}{f_Y(y)}dx=2y^2$$

$E[\cdot|Y=y]=h(y)\leftarrow$ 是关于y的函数

此时Y不一定是一个给定的集合，它可以是一个随机变量.

由Y是随机变量，那对于$E[X|Y]=h(Y)$来说，$E(X|Y)$也是一个随机变量

☆它不再是一个定值

三、对多个集合的条件期望值（decomposition 分解）

$(\Omega, \mathcal{F}, \mathbb{P})$: probability space

Let $\mathcal{D} = \{D_1, D_2, \cdots D_n\}$ satisfies

(i) $\mathcal{D}$ is a decomposition of Ω i.e.

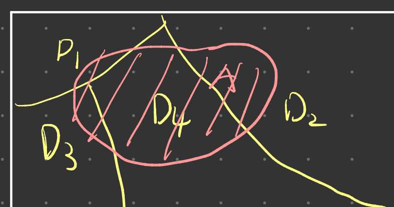
$$D_1 \uplus D_2 \uplus \cdots \uplus D_n = \Omega \quad \text{(disjoint)}$$

(ii) $D_i \in \mathcal{F}_i$, $\forall i$

(iii) $\mathbb{P}(D_i) > 0$, $\forall i$

Then for $A \in \mathcal{F}$, $\mathbb{P}(A \mid D_i)$ is well-defined for all i.

Definition : $\mathbb{P}(A \mid \mathcal{D}) = \sum_{i=1}^{n} \mathbb{P}(A \mid D_i) I_{D_i}$



$$\mathbb{P}(A \mid \mathcal{D})(\omega) = \sum_{i=1}^{n} \mathbb{P}(A \mid D_i) I_{D_i}$$

if $\omega \in D_1$: $\mathbb{P}(A \mid \mathcal{D})(\omega) = \mathbb{P}(A \mid D_1)$

例: Consider $([0,1], \mathcal{B}_1, m)$

$A = [\frac{1}{2}, \frac{3}{4}]$

$\mathcal{D} = \{[0, \frac{1}{3}], (\frac{1}{3}, \frac{2}{3}], [\frac{2}{3}, 1]\}$

Then $\mathbb{P}(A|\mathcal{D}) = \mathbb{P}(A|D_1) I_{D_1} + \mathbb{P}(A|D_2) I_{D_2} + \mathbb{P}(A|D_3) I_{D_3}$

$$= \frac{P(A, D_1)}{P(D_1)} + \frac{P(A, D_2)}{P(D_2)} + \frac{P(A, D_3)}{P(D_3)}$$

$$= \frac{0}{\frac{1}{3}} + \frac{\frac{2}{3} - \frac{1}{2}}{\frac{1}{3}} + \frac{\frac{3}{4} - \frac{2}{3}}{\frac{1}{3}}$$

$$= 0 + \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$$

Remark: (1) If $\mathcal{D} = \{\Omega\}$ ← 没有分割

$$\mathbb{P}(A|\mathcal{D}) = \mathbb{P}(A|\Omega) I_\Omega = \mathbb{P}(A)$$

(2) $\mathbb{P}(A|\mathcal{D})$ is a r.v.

If $\omega \in D_i$, $\mathbb{P}(A|\mathcal{D})(\omega) = \mathbb{P}(A|D_i)$

(3) If A, B are disjoint

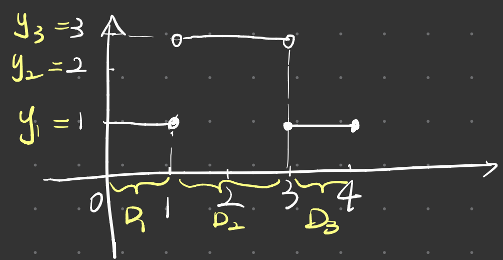
$$\mathbb{P}(A \cup B|\mathcal{D}) = \mathbb{P}(A|\mathcal{D}) + \mathbb{P}(B|\mathcal{D})$$

对于 discrete random variable (离散随机变量)构成的条件 decomposition

Let $Y = \sum_{i=1}^{n} y_i I_{D_i}$

where $\{y_1, y_2, \cdots, y_n\}$ = distinct constant

$\{D_1, D_2, \cdots, D_n\}$ = decomposition of Ω with $\mathbb{P}(D_i) > 0$, $\forall i$



Then $D_i = \{Y = y_i\}$

Definition: (1) $\mathbb{D}_Y = \{D_1, D_2, \cdots, D_n\}$ is called the decomposition induced by Y

(2) $\mathbb{P}(A|Y) := \mathbb{P}(A|\mathbb{D}_Y)$ is called the conditional probability of A w.r.t. the r.v. Y.

Moreover, we denote $\mathbb{P}(A|Y=y_i)$ the conditional probability $\mathbb{P}(A|D_i)$.

(3) Suppose $Y_1, Y_2, Y_3, \cdots, Y_n$ are sets of the form (*). We denote $\mathbb{D}_{Y_1, Y_2, \cdots Y_n}$ the decomposition induced by $Y_1, Y_2, \cdots, Y_n$.

Then $\mathbb{P}(A|\mathbb{D}_{Y_1, Y_2, \cdots Y_n}) = \mathbb{P}(A|Y_1, Y_2, \cdots, Y_n)$

例: (1) $\Omega = [0,10]$. $\mathbb{P} = \frac{1}{10} m$

$Y = 5 I_{[0,4]} + I_{[8,10]}$

$(\Rightarrow \mathcal{D}_Y = \{[0,4], (4,8), [8,10]\})$

$A = (3,5) \cup [7,9]$

$$\mathbb{P}(A|Y) = \mathbb{P}(A|\mathcal{D}_Y) = P(A|[0,4]) I_{[0,4]} + P(A|(4,8)) I_{(4,8)} + P(A|[8,10]) I_{[8,10]}$$

$$= \frac{P((3,4])}{P([0,4])} + \frac{P((4,5) \cup [7,8])}{P((4,8))} + \frac{P([8,9])}{P([8,10])}$$

$$= \frac{\frac{1}{10} \times 1}{\frac{1}{10} \times 4} + \frac{\frac{1}{10} \times (1+1)}{\frac{1}{10} \times 4} + \frac{\frac{1}{10} \times 1}{\frac{1}{10} \times 2}$$

$$= \frac{1}{4} + \frac{2}{4} + \frac{2}{4} = \frac{5}{4}$$

(2) X, Y: iid r.v.'s with $\mathbb{P}(X=1) = p$

$\mathbb{P}(X=0) = q = 1-p$

Let $A_k = \{X+Y=k\}$ for $k=0,1,2$

Compute $\mathbb{P}(A_k | Y)$

$$\mathbb{P}(A_k | Y=l) = \mathbb{P}(X+Y=k \mid Y=l) = \frac{\mathbb{P}(X+Y=k, Y=l)}{\mathbb{P}(Y=l)}$$

$$= \frac{\mathbb{P}(X=k-l, Y=l)}{\mathbb{P}(Y=l)}$$

$$= \mathbb{P}(X=k-l)$$

$$\Rightarrow \mathbb{P}(A_k|Y) = \mathbb{P}(A_k|Y=k) \cdot I_{\{Y=k\}} + \mathbb{P}(A_k|Y=k-1) \cdot I_{\{Y=k-1\}}$$

$$= \mathbb{P}(A_k|Y=0) I_{\{Y=0\}} + \mathbb{P}(A_k|Y=1) \cdot I_{\{Y=1\}}$$

当$k=0$时

$$\mathbb{P}(A_0|Y) = \mathbb{P}(X=0) I_{\{Y=0\}} + \mathbb{P}(X=-1) \cdot I_{\{Y=1\}}$$

$$= q \cdot I_{\{Y=0\}} = q \cdot (1-Y)$$

当$k=1$时

$$\mathbb{P}(A_1|Y) = \mathbb{P}(X=1) I_{\{Y=0\}} + \mathbb{P}(X=0) \cdot I_{\{Y=1\}}$$

$$= p(1-Y) + qY$$

$k=2$时

$$\mathbb{P}(A_2|Y) = \mathbb{P}(X=2) I_{\{Y=0\}} + \mathbb{P}(X=1) \cdot I_{\{Y=1\}}$$

$$= p \cdot Y$$

下定义$E[Y|\mathcal{D}]=?$

Let $X = \sum_{i=1}^{m} x_i I_{A_i}$ where $\{A_1, \cdots, A_n\}$ disjoint subset of Ω

$\{x_1, x_2, \cdots, x_m\}$ = distinct real numbers.

Then $E(X) = \sum_{i=1}^{m} x_i \mathbb{P}(A_i)$

Definition: The conditional expectation of X w.r.t. a decomposition $\mathcal{D} = \{D_1, D_2, \cdots, D_n\}$ is defined by $E(X|\mathcal{D}) = \sum_{i=1}^{m} x_i \mathbb{P}(A_i|\mathcal{D})$

Proposition: (1) $E[aX+bY|\mathcal{D}] = aE(X|\mathcal{D}) + bE(Y|\mathcal{D})$

(2) $E[c|\mathcal{D}] = c$, c is a constant

(3) $E[I_A|\mathcal{D}] = \mathbb{P}(A|\mathcal{D})$

(4) $E[E[X|\mathcal{D}]] = E[X]$

Remark: (1) $E[X|\mathcal{D}]$ is a r.v.

(2) $E[X|D] = \sum_{i=1}^{m} x_i \mathbb{P}(A_i|D)$

(3) $E[X|\mathcal{D}] = \sum_{i=1}^{m} x_i \mathbb{P}(A_i|\mathcal{D}) = \sum_{i=1}^{m} x_i \sum_{j=1}^{n} \mathbb{P}(A_i|D_j) I_{D_j}$

$$= \sum_{i=1}^{m}\sum_{j=1}^{n} x_i E(I_{A_i}|D_j) I_{D_j}$$

$$= \sum_{j=1}^{n} E\left(\sum_{i=1}^{m} x_i I_{A_i} | D_j\right) I_{D_j}$$

$$= \sum_{j=1}^{n} E(X|D_j) I_{D_j}$$

If $\omega \in D_i$, $E[X|\mathcal{D}](\omega) = E(X|D_i)$

X: discrete r.v. $\Rightarrow$ positive r.v.

$\Rightarrow X = X^+ - X^-$

例: $(\Omega, \mathcal{F}, \mathbb{P}) = ([0,1], \mathcal{B}_1, m)$

$\mathcal{D} = \{[0,\frac{1}{2}), [\frac{1}{2}, 1]\}$

$X = I_{[0,\frac{1}{3}]} + 2 I_{(\frac{1}{3},\frac{2}{3})} + 3 I_{(\frac{2}{3},1]}$

$$E[X|\mathcal{D}] = E[X|D_1] I_{D_1} + E[X|D_2] I_{D_2}$$

$$= \sum_i x_i P(x_i | [0,\tfrac{1}{2})) + \sum_i x_i P(x_i | [\tfrac{1}{2}, 1])$$

$$= 1 \times \frac{P([0,\frac{1}{3}])}{P([0,\frac{1}{2}])} + 2 \times \frac{P([\frac{1}{3},\frac{1}{2}])}{P([0,\frac{1}{2}])} + 2 \times \frac{P([\frac{1}{2},\frac{2}{3}])}{P([\frac{1}{2},1])} + 3 \times \frac{P([\frac{2}{3},1])}{P([\frac{1}{2},1])}$$

$$= 1 \times \frac{\frac{1}{3}}{\frac{1}{2}} + 2 \times \frac{\frac{1}{6}}{\frac{1}{2}} + 2 \times \frac{\frac{1}{6}}{\frac{1}{2}} + 3 \times \frac{\frac{1}{3}}{\frac{1}{2}}$$

$$= \frac{2}{3} + \frac{2}{3} + \frac{2}{3} + 2 = \frac{2}{3} \times 3 + 2 = 4$$

$$= \frac{4}{3} I_{D_1} + \frac{8}{3} I_{D_2}$$

当 $E[X|Y=y_i]$ 中 y_i 不再是可列个, 分得非常细, 该如何求 $E[X|Y]$?

$(\Omega, \mathcal{F}, \mathbb{P})$: probability space

$\mathcal{G} \subseteq \mathcal{F}$: σ-subalgebra

X: nonnegative r.v.

How to define $E[X|\mathcal{G}]$?

补充: X: $\mathcal{G}$-measurable if $\{X\in B\}\in\mathcal{G} \Longleftrightarrow \{X\le r\}\in\mathcal{G}$

重要!! ⟶

Definition: X对$\mathcal{G}$的条件期望: $E[X|\mathcal{G}]$ or $E[X|\mathcal{G}](\omega)$

满足:

(i) $E[X|\mathcal{G}]$ is $\mathcal{G}$-measurable ⟶ 是$\mathcal{G}$的函数

或说 $\{E[X|\mathcal{G}]\le r\}\in\mathcal{G}$, $\forall r\in\mathbb{R}$

(ii) For all $A\in\mathcal{G}$

$$\int_A E[X|\mathcal{G}]\,d\mathbb{P} = \int_A X\,d\mathbb{P}$$

$$E_A[E_{\mathcal{G}}(X)] = E_A(X)$$

例: (1) X: uniformly distributed on $([0,1],\mathcal{B}_1,m)$

$$\mathcal{G}=\sigma(\{0\le X<\tfrac{1}{2}\},\{\tfrac{1}{2}\le X\le 1\})$$

☆ 若X是$\mathcal{G}$-measurable 且$\mathcal{G}=\sigma(\{D_1,D_2,\cdots,D_n\})$ 则 $X=c_1 I_{D_1}+\cdots+c_n I_{D_n}$ ⟶

Then $E[X|\mathcal{G}]=c_1 I_{[0,\frac{1}{2})}+c_2 I_{[\frac{1}{2},1]}$ ($E[X|\mathcal{G}]$是$\mathcal{G}$-measurable)

Take $A=[0,\tfrac{1}{2})$, $\int_A E[X|\mathcal{G}]\,d\mathbb{P}=\int_A c_1\,d\mathbb{P}=\int_A X\,d\mathbb{P}$

$$c_1(\tfrac{1}{2}-0)=\int_0^{\frac{1}{2}} x\,dx=\tfrac{1}{8}$$

$$C_1 = \frac{1}{4}$$

Take $A = [\frac{1}{2}, 1]$, $\int_A E[X|\mathcal{G}]\,dP = \int_A C_2\,dP = \int_A X\,dP$

$$C_2(1-\tfrac{1}{2}) = \int_{\frac{1}{2}}^{1} x\,dx$$

$$\tfrac{1}{2}C_2 = \frac{3}{8}$$

$$C_2 = \frac{3}{4}$$

$$E[X|\mathcal{G}] = \frac{1}{4} I_{[0,\frac{1}{2})} + \frac{3}{4} I_{[\frac{1}{2},1]}$$

例2: (1) $([0,1], \mathcal{B}_1, m)$: probability space

X: r.v. with pdf $f(x) = 2x$

$\mathcal{G} = \sigma(\{0 \le X < \frac{1}{3}\}, \{\frac{1}{3} \le X < \frac{2}{3}\}, \{\frac{2}{3} \le X \le 1\})$

$$E[X|\mathcal{G}] = C_1 I_{[0,\frac{1}{3})} + C_2 I_{[\frac{1}{3},\frac{2}{3}]} + C_3 I_{[\frac{2}{3},1]}$$

Take $A = [0, \frac{1}{3})$

$$\int_A C_1\,dP = \int_A x\,dP$$

$$C_1 \int_0^{\frac{1}{3}} 2x\,dx = \int_0^{\frac{1}{3}} 2x^2\,dx$$

$$C_1 \cdot \frac{1}{9} = \frac{2}{3^4} \qquad C_1 = \frac{2}{3^4} \times 3^2 = \frac{2}{9}$$

Take $A = [\frac{1}{3}, \frac{2}{3}]$

$$\int_A C_2 \, dP = \int_A X \, dP$$

$$C_2 \int_{\frac{1}{3}}^{\frac{2}{3}} 2x \, dx = \int_{\frac{1}{3}}^{\frac{2}{3}} 2x^2 \, dx$$

$$C_2 = \frac{4}{9}$$

Take $A = [\frac{2}{3}, 1]$

$$\int_A C_3 \, dP = \int_A X \, dP$$

$$C_3 \int_{\frac{2}{3}}^{1} 2x \, dx = \int_{\frac{2}{3}}^{1} 2x^2 \, dx$$

$$C_3 = \frac{16}{45}$$

$$E[X|\mathcal{G}] = \frac{2}{9} I_{[0,\frac{1}{3})} + \frac{4}{9} I_{[\frac{1}{3},\frac{2}{3}]} + \frac{16}{45} I_{[\frac{2}{3},1]}$$

How about general case ?

Let X be a r.v. Then $X = X^+ - X^-$

Assumption: $\min\{E(X^+), E(X^-)\} < \infty$

The condition expectation of r.v. X w.r.t the σ-algebra $\mathcal{G}$ is defined by

$$E[X|\mathcal{G}] = E[X^+|\mathcal{G}] - E[X^-|\mathcal{G}]$$

当 $E[X^+|\mathcal{G}]$ 和 $E[X^-|\mathcal{G}]$ 都为 ∞，则 $E[X|\mathcal{G}]$ 可为任意数

下面通过条件期望去定条件概率：

仿照：$E(X) = \int_\Omega X \, d\mathbb{P}$

$E(I_A) = \int_A I_A \, d\mathbb{P} = \mathbb{P}(A)$

Definition: Let $B \in \mathcal{F}$, The conditional probability is defined by

$$\mathbb{P}(B|\mathcal{G}) = E(I_B|\mathcal{G})$$

Proposition:

(1) If $X \equiv C$, then $E(X|\mathcal{G}) = C$

(2) $X \leq Y$ $\mathbb{P}$-a.s. then

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

(3) $|E[X|\mathcal{G}]| \leq E[|X||\mathcal{G}]$

(4) $E[aX + bY|\mathcal{G}] = aE[X|\mathcal{G}] + bE[Y|\mathcal{G}]$

(5) $\mathcal{F}^* = \{\Omega, \emptyset\}$:= trivial σ-algebra.

Then $E[X|\mathcal{F}^*] = E[X]$

(6) If X is $\mathcal{G}$-measurable

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

(7) $E[E[X|\mathcal{G}]] = E[X]$

用第2条 $\int_A E[X|\mathcal{G}]\,dP = \int_A X\,dP$

(8) (Tower property)

If $\mathcal{G}_1 \subseteq \mathcal{G}_2$

$$E[E[X|\mathcal{G}_1]|\mathcal{G}_2] = E[X|\mathcal{G}_1] = E[E[X|\mathcal{G}_2]|\mathcal{G}_1]$$

(9) Y: $\mathcal{G}$-measurable, $E|X|<\infty$, $E|XY|<\infty$.

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

定理: (X_n): sequence of r.v.'s

(1) (Conditional Dominated Convergence Theorem)

If $|X_n| \le Y$, $EY<\infty$ and $X_n \to X$ P-a.s.

$E[X_n|\mathcal{G}] \to E[X|\mathcal{G}]$ P-a.s.

(2) (Conditional Monotone Convergence Theorem)

(i) If $X_n \geq Y$, $E[Y] > -\infty$ and $X_n \uparrow X$ P-a.s.

then $E[X_n|\mathcal{G}] \uparrow$ 죄 $E[X|\mathcal{G}]$ P-a.s.

(ii) If $X \leq Y$, $E[Y] < \infty$ and $X_n \downarrow X$ P-a.s.

then $E[X_n|\mathcal{G}] \downarrow$ 죄 $E[X|\mathcal{G}]$ P-a.s.

(3) (Conditional Fatou's Lemma).

(i) If $X_n \geq Y$, $E[Y] > -\infty$, then

$$E[\liminf_n X_n|\mathcal{G}] \leq \liminf_n E[X_n|\mathcal{G}]$$

(ii) If $X_n \leq Y$, $E[Y] < \infty$, then

$$\limsup_n E[X_n|\mathcal{G}] \leq E[\limsup_n X_n|\mathcal{G}]$$

(4) If $X_n \geq 0$ then

$$E[\sum_n X_n|\mathcal{G}] = \sum_n E[X_n|\mathcal{G}]$$

(5) Conditional Jesen's inequality

If φ is a convex function on $\mathbb{R}$. $E[X] < \infty$, $E[\varphi(X)] < \infty$

then $\varphi(E[X|\mathcal{G}]) \leq E[\varphi(X)|\mathcal{G}]$

思路：

概率论

$P(A|B) \xrightarrow{\text{把A变为r.v. X}} P(X|B),\ E(X|B)$

$P(A|B) \xrightarrow{\text{把B换成 } \mathcal{D}=\{D_1,\cdots,D_n\} \text{ decomposition}} \mathbb{P}(A|\mathcal{D}) \xrightarrow{\text{定义出}} E[X|\mathcal{D}]$

$P(A|B) \xrightarrow{\text{把A, B都换成 discrete r.v. X和Y}} P(X=x|Y=y) \longrightarrow E(X|Y=y)$

$P(X=x|Y=y) \xrightarrow{\text{continuous r.v. X和Y}} f(X=x|Y=y) \longrightarrow E(X|Y=y)$

$E(X|Y=y) \longrightarrow E(X|Y)=h(Y)$

也是r.v.

$\mathbb{P}(A|\mathcal{D}) \xrightarrow{\mathcal{D}\text{是由discrete r.v. Y生成的}} \mathbb{P}(A|Y)=\mathbb{P}(A|\mathcal{D}_Y) \xrightarrow{\text{定义出}} E[X|Y]$

此处A无需是r.v.

$\mathbb{P}(A|Y) \xrightarrow{\text{当Y是continuous时 } \mathcal{D} \text{ 不再好分，此处先引入条件期望}} E[X|\mathcal{G}]$ $\mathcal{G}$: σ-algebra

$E[X|\mathcal{G}] \xrightarrow{\text{由期望推出概率}} P(X|\mathcal{G})$

$E(X|Y)=h(Y) \xleftrightarrow{\text{一样的结果}} E[X|Y]$