確率数理要論5

条件付き期待値

これまで独立なr.v.の最大値や平均を考えてきたが、
非独立なr.v.の組を考える上で条件付期待や条件付き確率は重要である。

まず有限分割の場合から始める。

$(\Omega, \mathcal{F}, P)$：確率空間
$B \in \mathcal{F}$ を $P(B) > 0$ をみたす事象とする。
$\forall A \in \mathcal{F}$ に対し、Bで条件付けたAの条件付き確率を

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

とする。これが確率測度になっていることはすぐわかる。
（つまり、$A \in \mathcal{F} \mapsto P(A|B)$ が確率測度）
ここで、$B^c \in \mathcal{F}$ も $P(B^c) > 0$ をみたすとすると、

$$P(\cdot|B), \quad P(\cdot|B^c)$$

がそれぞれ定まる。つまり、$\forall A \in \mathcal{F}$ に対し、確率変数

$$\Omega \ni \omega \longmapsto \begin{cases} P(A|B) & (\omega \in B) \\ P(A|B^c) & (\omega \in B^c) \end{cases}$$

$$= P(A|B) 1_B(\omega) + P(A|B^c) 1_{B^c}(\omega)$$

$$=: P(A|\{B, B^c\})(\omega)$$

が定まる。
また、可積分な X に対して、X の $P(\cdot|B)$ による期待値を条件付き期待値といい、

$$E[X|B] \qquad \left(つまり、E[X|B] = \frac{1}{P(B)} \int_B X \, dP(\omega)\right)$$

と書く。これについても

$\int 1_B(\omega) X(\omega) dP(\omega)$ という意味

$$E[X|\{B, B^c\}](\omega) = E[X|B] 1_B(\omega) + E[X|B^c] 1_{B^c}(\omega)$$

なる確率変数が定まる。

これを有限分割 $\{B_i\}_{i=1}^n$ $\left(B_i \cap B_j = \phi\ (i \neq j),\ \bigcup_{i=1}^n B_i = \Omega,\ P(B_i) > 0\right)$
に拡張することで、

$$P(A \mid \{B_i\}_{i=1}^n)(\omega) := \sum_{i=1}^n P(A \mid B_i)\,\mathbb{1}_{B_i}(\omega)$$

$$E[X \mid \{B_i\}_{i=1}^n](\omega) := \sum_{i=1}^n E[X \mid B_i]\,\mathbb{1}_{B_i}(\omega)$$

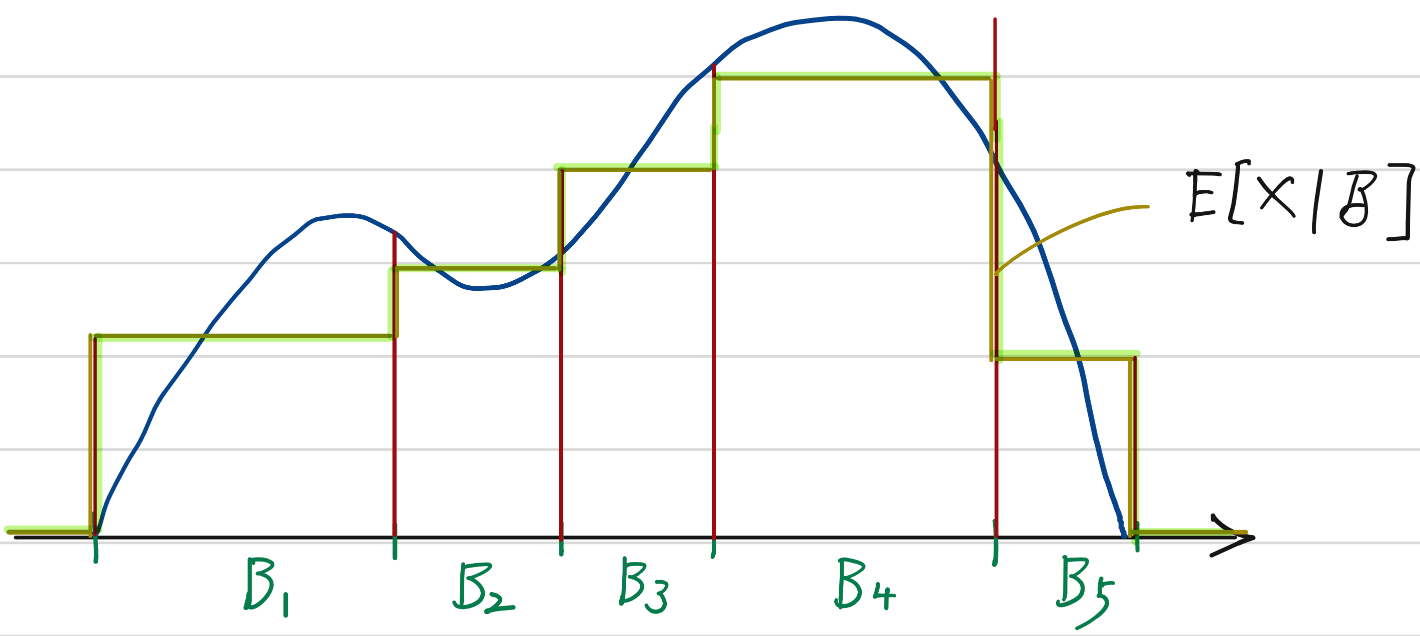
なる確率変数が定まる。

$\{B_i\}_{i=1}^n$ を与えることと、$\{B_i\}_{i=1}^n$ で生成される σ-加法族 $\mathcal{B}$ を与えることは
同値なので

$$P(A \mid \mathcal{B}),\quad E[X \mid \mathcal{B}]$$

と書く。

$X = \mathbb{1}_A$ なら、$E[X \mid \mathcal{B}] = P(A \mid \mathcal{B})$ となることは期待値の定義から
すぐ確認できる。



Thm

1. $E[X|\mathcal{B}]$ は $\mathcal{B}$-可測

2. $^\forall G \in \mathcal{B}$ に対し、
$$\int_G E[X|\mathcal{B}]\,dP(\omega) = \int_G X\,dP(\omega)$$
//

※ $E[X|\mathcal{B}]$ が $\mathcal{F}$-可測なのはすぐ分かるが、1 はより強いことを言っている。
($\mathcal{B}$ は $\mathcal{F}$ の部分σ-加法族)

※ $G \in \mathcal{B}$ ということは、ある $I \subset \{1, 2, \dots, n\}$ が存在して、($I = \phi$ も許す)
$G = \bigcup_{i \in I} B_i$ と書けることに注意。

(Proof)

1. $C \in \mathcal{B}(\mathbb{R})$ に対し、$I_C = \{i \mid E[X|B_i] \in C\}$ とおけば
$E[X|\mathcal{B}]^{-1}(C) = \bigcup_{i \in I_C} B_i$ であるが。

$\bigcup_{i \in I_C} B_i \in \mathcal{B}$ であるので、$E[X|\mathcal{B}]^{-1}(C) \in \mathcal{B}$ である。

2. $^\forall G \in \mathcal{B}$ は、ある $I \subset \{1, \dots, n\}$ を用いて、$G = \bigcup_{i \in I} B_i$ と書ける。
よって、
$$\int_G E[X|\mathcal{B}]\,dP(\omega) = \sum_{i \in I} \int_{B_i} E[X|\mathcal{B}]\,dP(\omega)$$
$$= \sum_{i \in I} \int_{B_i} \sum_{j=1}^{n} E[X|B_j]\, \mathbb{1}_{B_j}(\omega)\,dP(\omega)$$
$$= \sum_{i \in I} \sum_{j=1}^{n} E[X|B_j] \int_{B_i} \mathbb{1}_{B_j}(\omega)\,dP(\omega)$$
$$= \sum_{i \in I} \sum_{j=1}^{n} E[X|B_j]\, P(B_i)\, \delta_{ij}$$
$$= \sum_{i \in I} P(B_i)\, E[X|B_i]$$
$$= \sum_{i \in I} P(B_i) \frac{1}{P(B_i)} \int_{B_i} X\,dP(\omega)$$
$$= \sum_{i \in I} \int \mathbb{1}_{B_i} X\,dP(\omega) = \int \left(\sum_{i \in I} \mathbb{1}_{B_i}\right) X\,dP(\omega)$$
$$= \int \mathbb{1}_{\bigcup_{i \in I} B_i} X\,dP(\omega) = \int \mathbb{1}_G X\,dP(\omega) = \int_G X\,dP(\omega)$$
//

これを一般化しよう.

$\mathcal{G}$ を $\mathcal{F}$ の部分 σ-加法族とする.

Remind: $\sigma(Y)$ は Y を可測にする最小の σ-加法族.

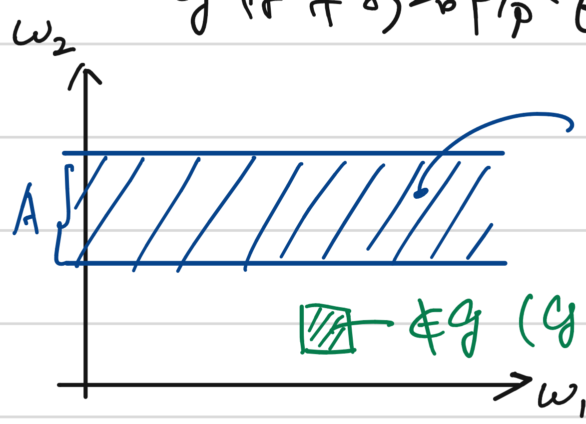
例. $\mathcal{F} = \mathcal{B}(\mathbb{R}^2)$

$\omega = (\omega_1, \omega_2) \in \mathbb{R}^2$, $X(\omega) = \omega_1$, $Y(\omega) = \omega_2$ のとき,

$$\mathcal{G} = \sigma(Y) = \{Y^{-1}(A) \mid A \in \mathcal{B}(\mathbb{R})\} = \{\Omega \times A \mid A \in \mathcal{B}(\mathbb{R})\}$$

このとき,

$\mathcal{G}$ は $\mathcal{F}$ の部分 σ-加法族



$\Omega \times A \in \mathcal{G}$ (含まれる)

$\notin \mathcal{G}$ ($\mathcal{G}$ には含まれない)

//

Def (条件付き期待値)

X: 可積分な r.v. on $(\Omega, \mathcal{F}, P)$

X の $\mathcal{G}$ で条件付けた条件付き期待値 $E[X \mid \mathcal{G}]$ とは,

1. $\mathcal{G}$-可測である.
2. $^\forall G \in \mathcal{G}$ に対し,

$$\int_G E[X \mid \mathcal{G}] \, dP(\omega) = \int_G X \, dP(\omega)$$

を満たす確率変数として定義される. //

※ G は $\mathcal{G}$ の元であって $\mathcal{F}$ の元ではないことに注意!

特に, $A \in \mathcal{F}$ に対して, $X = \mathbb{1}_A$ を代入して,

$$P(A \mid \mathcal{G}) = E[\mathbb{1}_A \mid \mathcal{G}]$$

を A の ($\mathcal{G}$ で条件付けた) 条件付き確率と言う.

では, そのような確率変数は存在するであろうか?

(1.2. より $P|_{\mathcal{G}}$ − a.s. での一意性は言える. ($P|_{\mathcal{G}}$ は $\mathcal{G}$ に制限した P)

なぜならもし Z_1, Z_2 が 1.2. をみたしつつ, $P(Z_1 > Z_2 + \varepsilon) > \varepsilon$ なら,

$G = \{Z_1 > Z_2 + \varepsilon\}$ とすると, $\underset{\displaystyle \int_G X \, dP}{\underset{\shortparallel}{\int_G Z_1 \, dP}} \geq \int_G Z_2 \, dP + \varepsilon P(G) \geq \int_G X \, dP + \varepsilon^2$ で矛盾.)

(∵ 2より)

Thm (Radon-Nikodymの定理)

$(\Omega, \mathcal{F}, P)$：確率空間
ν：$(\Omega, \mathcal{F})$上の非負かつ有界な測度でPに対して絶対連続、
つまり、$P(A)=0$ なら $\nu(A)=0$ をみたすものとする（$\nu \ll P$ と書く）
このとき、P-可積分な r.v. X が存在して、

$$\nu(A) = \int_A X \, dP \qquad (\forall A \in \mathcal{F})$$

が成り立つ. しかも、XはP-a.s.で一意に決まる. //

X を Radon-Nikodym微分と言い $X = \frac{d\nu}{dP}$ と書く.

（証明は、例えば Resnik: A probability Path を参照。このノートの最後にも証明をのせておいた。）

Thm (条件付き期待値の存在)

Xが可積分なら条件1,2をみたす確率変数 $E[X|\mathcal{G}]$ が存在する.
しかも、2つのr.v. Z_1, Z_2 がともに条件1,2をみたすなら

$$Z_1 = Z_2 \qquad (P|_\mathcal{G} - a.s.)$$

が成り立つ.　← Pを$\mathcal{G}$に制限した測度
この意味で条件付き期待値は$P|_\mathcal{G}$-a.s.で一意. //

(Proof)

$X \geq 0$ の場合を考える.
$E[|X|] < \infty$ より、

$$\nu(A) = \int_A X \, dP \qquad (\forall A \in \mathcal{G})$$

とすれば、νは$\mathcal{G}$上の非負有界な測度である.
しかも、$P(A)=0$ $(A \in \mathcal{G})$ なら $\nu(A)=0$ もすぐわかる.
つまり、$\nu \ll P|_\mathcal{G}$.
よって、Radon-Nikodymの定理より$\mathcal{G}$-可測なr.v. Zが$P|_\mathcal{G}$-a.s.で一意に存在し、

$$\nu(A) = \int_A Z \, dP|_\mathcal{G}(\omega) = \int_A Z \, dP(\omega) \qquad (\because P = P|_\mathcal{G} \text{ on } \mathcal{G},\ Z\text{は}\mathcal{G}\text{-可測, 積分の定義})$$

を満たす.

この Z を $E[X|\mathcal{G}]$ と書けば定理が示される.
一般の X に対しては、$X = X^+ - X^-$ と $X^+ \geq 0, X^- \geq 0$ を用いて分解し、それぞれに対し同じ議論を適用すれば良い. //

Thm (条件付き期待値の性質)

(1) (線形性)

$a, b \in \mathbb{R}$ に対し、

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

(測度0で除外される集合は a, b に依存)

(2) (単調性)

$X \geq 0$ (a.s.), X:可積分 なら、

$$E[X|\mathcal{G}] \geq 0 \quad (a.s.)$$

特に、$X \geq Y$ (a.s.), X, Y:可積分なら

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

である.

(3) (絶対値の不等式)

$$|E[X|\mathcal{G}]| \leq E[|X||\mathcal{G}] \quad (a.s.)$$

(4) (単調収束定理)

$E[|X|] < \infty$ かつ、$0 \leq X_1 \leq X_2 \leq \cdots \leq X$ で、$\lim_{n\to\infty} X_n = X$ (a.s.) のとき、

$$\lim_{n\to\infty} E[X_n|\mathcal{G}] = E[X|\mathcal{G}] \quad (a.s.)$$

(これより、条件付き期待値の Fatou の補題と優収束定理も成り立つ)

☆ (5) Y が $\mathcal{G}$-可測で X と XY が可積分なら

$$E[XY|\mathcal{G}] = YE[X|\mathcal{G}] \quad (a.s.)$$

特に $E[Y|\mathcal{G}] = Y$ (a.s.)

☆ (6) $\mathcal{G}'$ が $\mathcal{G}$ の部分σ-加法族なら

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

(7) $\sigma(X)$: X の生成するσ-加法族 $(\sigma(X) = \{X^{-1}(A) \mid A \in \mathcal{B}(\mathbb{R})\})$

$\sigma(X)$ が $\mathcal{G}$ と独立なら

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

//

(Proof)

(3)
$$\begin{aligned} |E[X|\mathcal{G}]| &= |E[X^+|\mathcal{G}] - E[X^-|\mathcal{G}]| \quad (X = X^+ - X^- \text{と分解}) \\ &\leqq |E[X^+|\mathcal{G}]| + |E[X^-|\mathcal{G}]| \\ &= E[X^+|\mathcal{G}] + E[X^-|\mathcal{G}] \\ &= E[X^+ + X^-|\mathcal{G}] = E[|X||\mathcal{G}] \end{aligned}$$

(4) 単調性より $(E[X_n|\mathcal{G}])_{n=1}^{\infty}$ も a.s. で単調非減少

$E[X_n|\mathcal{G}] \leqq E[X|\mathcal{G}]$ (a.s.) より、ある $\mathcal{G}$-可測な r.v. Z が存在して、

$$Z = \lim_{n\to\infty} E[X_n|\mathcal{G}] \quad (a.s.)$$

とできる。 (∵ 有界な単調列は収束する)

$\forall A \in \mathcal{G}$ に対して、

$$\begin{aligned} \int_A Z\,dP &= \int_A \lim_{n\to\infty} E[X_n|\mathcal{G}]\,dP \\ &= \lim_{n\to\infty} \int_A E[X_n|\mathcal{G}]\,dP \quad (\text{単調収束定理}) \\ &= \lim_{n\to\infty} \int_A X_n\,dP \quad (\text{条件付き期待値の定義}) \\ &= \int_A \lim_{n\to\infty} X_n\,dP \quad (\text{再び単調収束定理}) \\ &= \int_A X\,dP \end{aligned}$$

よって、

$$\lim_{n\to\infty} E[X_n|\mathcal{G}] = Z = E[X|\mathcal{G}] \quad (a.s.)$$

(∵ 条件付き期待値の定義と一意性より) //

(5) ある $A \in \mathcal{G}$ を用いて $Y = \mathbb{1}_A$ と書けるなら
$^\forall B \in \mathcal{G}$ に対して

$$\int_B E[XY \mid \mathcal{G}]\, dP = \int_B XY\, dP$$
$$= \int_{B \cap A} X\, dP$$
$$= \int_{B \cap A} E[X \mid \mathcal{G}]\, dP \quad (\because B \cap A \in \mathcal{G})$$
$$= \int_B \underbrace{\mathbb{1}_A E[X \mid \mathcal{G}]}_{Y E[X \mid \mathcal{G}]}\, dP$$

Y が $\mathcal{G}$-可測なので、$Y E[X \mid \mathcal{G}]$ も $\mathcal{G}$-可測。よって条件付き期待値の一意性より $E[XY \mid \mathcal{G}] = Y E[X \mid \mathcal{G}]$ (a.s.) である。

より一般に $X \geqq 0$, $Y \geqq 0$ のとき、
Y はある単関数列 $Y_n = \sum_{i=1}^{k_n} a_i \mathbb{1}_{A_i}$ の単調収束極限とできることから、単調収束定理から示せる。

$X \geqq 0$, $Y \geqq 0$ でないときは、$X = X^+ - X^-$, $Y = Y^+ - Y^-$ と分解すれば良い。

(6) $^\forall A \in \mathcal{G}'$ に対し、$A \in \mathcal{G}$ でもあるので、

$$\int_A E[E[X \mid \mathcal{G}] \mid \mathcal{G}']\, dP = \int_A E[X \mid \mathcal{G}]\, dP \quad (\because A \in \mathcal{G}')$$
$$= \int_A X\, dP \quad (\because A \in \mathcal{G})$$
$$= \int_A E[X \mid \mathcal{G}']\, dP$$

(再び $A \in \mathcal{G}'$ と条件付き期待値の定義より)

よって、$E[E[X \mid \mathcal{G}] \mid \mathcal{G}'] = E[X \mid \mathcal{G}']$ (a.s.) ($\because$ 一意性)

//

Thm（Jensenの不等式）

$\phi: \mathbb{R} \to \mathbb{R}$ を凸関数とする.

X と $\phi(X)$ がともに可積分なら、

$$\phi(E[X|\mathcal{G}]) \leqq E[\phi(X)|\mathcal{G}] \quad \text{(a.s)}$$

（略証）

ϕ の $E[X|\mathcal{G}]$ における劣微分を $t \in \partial\phi(E[X|\mathcal{G}])$ としてとってくる

（t は $\mathcal{G}$-可測 なもの.）

${}^\forall x$ で $\phi(x) \geqq \phi(E[X|\mathcal{G}](\omega)) + \langle t(\omega), x - E[X|\mathcal{G}] \rangle$ である.

両辺 $E[\cdot|\mathcal{G}]$ をとれば示せる.

（本当は t の有界性が問題になるので. $X_n = X \mathbb{1}_{\{|X| \leqq n\}}$ として

X_n について示し、$n \to \infty$ とする.）

//

- 正則な条件付き確率

上の条件付き確率は各 $A \in \mathcal{F}$ ごとに $P(A|\mathcal{G})$ が a.s で

一意に決まった.

よって互いに排反な $(A_i)_{i=1}^{\infty} \in \mathcal{F}$ に対し（これは1つ固定する）

$$\sum_{i=1}^{\infty} P(A_i|\mathcal{G})(\omega) = P\left(\bigcup_{i=1}^{\infty} A_i \middle| \mathcal{G}\right)(\omega)$$

も a.s. で成り立たせることができる.

（∵ 左辺で定義される r.v. が単調収束定理より、条件付き確率の

条件を満たすため.

$${}^\forall B \in \mathcal{G}, \quad \int_B \sum_{i=1}^{\infty} P(A_i|\mathcal{G})\,dP = \sum_{i=1}^{\infty} \int_B P(A_i|\mathcal{G})\,dP$$

$$= \sum_{i=1}^{\infty} P(A_i) = P\left(\bigcup_{i=1}^{\infty} A_i\right)$$

$$= \int_B P\left(\bigcup_{i=1}^{\infty} A_i \middle| \mathcal{G}\right) dP$$

）

しかし. $(A_i)_{i=1}^{\infty}$ の選び方として、非可算無限個の可能性がありえるので、各固定された ω に対して. $(A_i)_{i=1}^{\infty}$ の選び方によらず一様にこの等式を成り立たせられるかは自明ではない.

つまり. $P(\cdot|\mathcal{G})(\omega)$ が確率1で確率測度になるとは限らない.

ω に関して a.s.

しかし、Ωが完備可分距離空間かつ$\mathcal{F}$がそのBorel集合族なら、次をみたす正則条件付き確率 $\mu : \Omega \times \mathcal{F} \longrightarrow [0,1]$ が存在する。

(1) $^\forall A \in \mathcal{F}$に対し、$\omega \longmapsto \mu(\omega, A)$ が $P(A|\mathcal{G})$ と a.s. で一致

(2) a.s. のωに対し、$A \longmapsto \mu(\omega, A)$ が $(\Omega, \mathcal{F})$ 上の確率測度

(証明は伊藤清「確率論」や Dudley「Real analysis and Probability」を参照)

XとYの同時分布が密度関数$f(x,y)$を持つとき、

$$P(Y \in A \mid \sigma(X))(\omega) = \frac{\int_A f(X(\omega), y)\, dy}{f(X(\omega))} \qquad \left(f(x) = \int f(x,y)\,dy \right)$$

である。

つまり、条件付き分布の密度は $f(y|x) = \dfrac{f(x,y)}{f(x)}$ で与えられる。

[Radon-Nikodymの定理の証明]

νは有界かつ非負なので.

$$Q(A) = \frac{\nu(A)}{\nu(\Omega)} \qquad (A \in \mathcal{F})$$

は確率測度になる. また, $Q \ll P$ であることもすぐわかる.

ここで.

$$P^* = \frac{P+Q}{2}$$

とする.

$L^2(P^*)$を P^*による2乗可積分な関数のなすヒルベルト空間とする.

$$L^2(P^*) = \{ Y \mid \int Y^2 dP^* < \infty \},$$

$$\langle Y, Y' \rangle = \int YY' dP^*.$$

そこで

$$L(Y) = \int Y dQ \qquad (Y \in L^2(P^*))$$

と線形汎関数を定義する.

これは有界線形汎関数であることが以下のようにしてわかる.

$$|L(Y)| \leq \int |Y| dQ \leq \int |Y| dQ + \int |Y| dP$$

$$= 2\int |Y| dP^* \underset{\substack{\uparrow \\ \text{Jensen}}}{\leq} 2\sqrt{\int |Y|^2 dP^*} = 2\|Y\|_{L^2(P^*)}$$

よって. Lは有界線形なので. $^\exists Z \in L^2(P^*)$で

$$L(Y) = \int YZ\, dP^* = \langle Y, Z \rangle$$

と表すことができる.

$$\underset{\substack{\| \\ \int Y dQ}}{\int YZ\, dP^*} = \frac{1}{2}\int YZ\, dP + \frac{1}{2}\int YZ\, dQ$$

より. $\int Y(1 - \frac{Z}{2}) dQ = \int \frac{YZ}{2} dP$ を得る. $(^\forall Y \in L^2(P^*))$

$Y = 1_A$ とすると、

$$L(Y) = \int Y \, dQ = Q(A) = \int 1_A Z \, dP^* = \int_A Z \, dP^*$$

である。両辺 $P^*(A)^{-1}$ 倍することで、

$$0 \le \frac{Q(A)}{P^*(A)} \le \frac{\int_A Z \, dP^*}{P^*(A)} \le \frac{\int_A Z \, dP^*}{Q(A)/2} = 2$$

$\left(\because P^* = \frac{Q+P}{2}\right)$

であることがわかる。

$\forall A$ に対して、

$$0 \le \int_A Z \, dP^* \le 2 P^*(A) = \int_A 2 \, dP^*$$

であることから、

$$0 \le Z \le 2 \quad (P^* \text{ a.s.})$$

である。

さらに、$Y = 1_{\{Z=2\}}$ とすると、

$$\int_{\{Z=2\}} \left(1 - \frac{Z}{2}\right) dQ = \int_{\{Z=2\}} \frac{Z}{2} \, dP$$

であるので $0 = P(Z=2)$ を得る。($\because$ 左辺 $=0$, 右辺 $= P(Z=2)$)

$Q \ll P$ でもあるので、$P^*(Z=2) = 0$ を得る。

つまり $0 \le Z < 2$ (P^*-a.s.) である。

今

$$Y = \left(\frac{Z}{2}\right)^n 1_A \quad (A \in \mathcal{F})$$

とすると、$Y \in L^2(P^*)$ であり $0 \le Y < 1$ (P^* a.s.) である。

また、

$$\int_A \left(\frac{Z}{2}\right)^n \left(1 - \frac{Z}{2}\right) dQ = \int_A \left(\frac{Z}{2}\right)^n \frac{Z}{2} \, dP$$

より、両辺 $\sum_{n=0}^{N}$ することに

$$\int_A \left(1 - \frac{Z}{2}\right)^{N+1} dQ = \int_A \frac{Z}{2} \sum_{j=0}^{N} \left(\frac{Z}{2}\right)^j dP$$

となる。今 $N \to \infty$ で $\left(1-\frac{Z}{2}\right)^{N+1} \nearrow 1$ (P^*-a.s.) なので

左辺 $\to \int_A dQ = Q(A)$

右辺 $\to \int_A \frac{Z/2}{1-Z/2} dP = \int_A \frac{Z}{2-Z} dP$

である。あとは $X = \frac{Z}{2-Z}$ とすれば O.K.

//