§7. 条件付き確率と条件付き期待値

§6 において扱った Fisher 計量に関する性質を更に調べるための準備として, 条件付き確率および条件付き期待値について述べよう.
まず, $(\Omega, \mathcal{F}, P)$ を確率空間とし, $B \in \mathcal{F}$, $P(B) > 0$ とする. このとき, 可測空間 $(\Omega, \mathcal{F})$ 上の確率測度 P_B を

$$P_B(A) = \frac{P(A \cap B)}{P(B)} \quad (A \in \mathcal{F})$$

により定めることができる. $P_B(A)$ を B に関する A の条件付き確率という. $P_B(A)$ は $P(A|B)$ とも表す.
X を $(\Omega, \mathcal{F}, P)$ 上の P 可積分な確率変数とすると, X を $(\Omega, \mathcal{F}, P_B)$ 上の P_B 可積分な確率変数とみなすことができる. P_B に関する X の期待値を $E[X|B]$ と表す. すなわち,

$$\begin{aligned} E[X|B] &= \int_\Omega X(\omega) P_B(d\omega) \\ &= \frac{1}{P(B)} \int_B X(\omega) P(d\omega) \end{aligned}$$

である. $E[X|B]$ を B に関する X の条件付き平均値または条件付き期待値という.
特に, $A \in \mathcal{F}$ とすると, A の定義関数 χ_A は $(\Omega, \mathcal{F}, P)$ 上の P 可積分な確率変数で,

$$E[\chi_A|B] = P(A|B)$$

がなりたつ.
次に, 上に現れた B を同時にいくつも考え, $\mathcal{F}$ の部分 σ 加法族に関する条件付き期待値および条件付き確率を定めよう.
X を $(\Omega, \mathcal{F}, P)$ 上の P 可積分な確率変数, $\mathcal{G}$ を $\mathcal{F}$ の部分 σ 加法族とし, $B \in \mathcal{G}$ に対して

$$Q(B) = E[X, B] = \int_B X(\omega) P(d\omega)$$

とおく.
このとき, Q は $(\Omega, \mathcal{G})$ 上の有限符号付き測度となる. また, P を $\mathcal{G}$ に制限して得られる $(\Omega, \mathcal{G})$ 上の確率測度も P と表すことにすると, Q は P に関して絶対連続となる.
よって, Radon-Nikodym の定理より, $(\Omega, \mathcal{G}, P)$ 上の P 可積分な確率変数 Y が存在し, 任意の $B \in \mathcal{G}$ に対して

$$Q(B) = \int_B Y(\omega) P(d\omega)$$

がなりたつ. この Y を $E[X|\mathcal{G}]$ と表し, $\mathcal{G}$ に関する X の条件付き平均値または条件付き期待値という.
なお, Y および Y' がともに $\mathcal{G}$ に関する X の条件付き期待値ならば,

$$Y = Y' \quad \text{a.s.}$$

となるから, 確率 1 で等しい確率変数を同一視すれば, 条件付き期待値の存在は一意的である. 以下では, このような同一視を行うことにする.
更に, $A \in \mathcal{F}$ のとき, $E[\chi_A|\mathcal{G}]$ を $P(A|\mathcal{G})$ と表し, $\mathcal{G}$ に関する A の条件付き確率という.

例　X を $(\Omega, \mathcal{F}, P)$ 上の P 可積分な確率変数とし, $\mathcal{F}$ の部分 σ 加法族 $\mathcal{G}$ として自明な σ 加法族を選んでおこう. すなわち,

$$\mathcal{G} = \{\emptyset, \Omega\}$$

である.
このとき,

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

であることはほとんど明らかである.

例　X を $(\Omega, \mathcal{F}, P)$ 上の P 可積分な確率変数, $\{B_i\}_{1\le i\le n}$ を Ω の $\mathcal{F}$ 可測な有限分割とする. すなわち,

$$B_i \in \mathcal{F},\ P(B_i) > 0\ (i = 1, 2, \ldots, n),\ B_i \cap B_j = \emptyset\ (i \ne j),\ \Omega = \bigcup_{i=1}^{n} B_i$$

である. $\mathcal{F}$ の部分 σ 加法族 $\mathcal{G}$ として $\{B_i\}_{1\le i\le n}$ を含む最小の σ 加法族を選んでおこう. これを

$$\mathcal{G} = \sigma(\{B_i\}_{1\le i\le n})$$

と表すことが多い.
このとき,

$$E[X|\mathcal{G}] = \sum_{i=1}^{n} E[X|B_i]\chi_{B_i}$$

である. 実際, 上の式の右辺は $(\Omega, \mathcal{G}, P)$ 上の確率変数で, $j = 1, 2, \ldots, n$ とすると,

$$\int_{B_j} X(\omega)P(d\omega) = \int_{B_j} \sum_{i=1}^{n} E[X|B_i]\chi_{B_i} P(d\omega) = E[X|B_j]P(B_j)$$

がなりたつからである.
更に, $A \in \mathcal{F}$ とすると,

$$P(A|\mathcal{G}) = \sum_{i=1}^{n} P(A|B_i)\chi_{B_i}$$

となるから, これは最初に述べた条件付き確率の自然な一般化とみなすことができる.

条件付き期待値の基本的な性質について述べよう.

定理　X, Y を $(\Omega, \mathcal{F}, P)$ 上の P 可積分な確率変数, $\mathcal{G}$ を $\mathcal{F}$ の部分 σ 加法族とする. このとき, 次の (1)〜(4) がなりたつ.
(1) $a, b \in \mathbf{R}$ とすると,

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

(2) $X \ge 0$ ならば,

$$E[X|\mathcal{G}] \ge 0.$$

(3) X が $\mathcal{G}$ 可測で, XY が P 可積分ならば,

$$E[XY|\mathcal{G}] = XE[Y|\mathcal{G}].$$

特に,

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

(4) $\mathcal{H}$ を $\mathcal{G}$ の部分 σ 加法族とすると,

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

特に,

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

証明　(1), (2): ほとんど明らか.
(3): 任意の $B \in \mathcal{G}$ に対して

$$E[XY, B] = E[XE[Y|\mathcal{G}], B]$$

がなりたつことを示せばよい.
まず, $A \in \mathcal{G}$ を用いて $X = \chi_A$ と表される場合,

$$\begin{aligned} E[\chi_A E[Y|\mathcal{G}], B] &= E[E[Y|\mathcal{G}], A \cap B] \\ &= E[Y, A \cap B] \\ &= E[\chi_A Y, B]. \end{aligned}$$

X が一般の場合は極限操作を行えばよい.
(4): $B \in \mathcal{H}$ とすると,

$$\begin{aligned} E[E[E[X|\mathcal{G}]|\mathcal{H}], B] &= E[E[X|\mathcal{G}], B] \\ &= E[X, B]. \end{aligned}$$

□

次の Jensen の不等式は条件付き期待値に関する重要な不等式である.

Jensen の不等式　X を $(\Omega, \mathcal{F}, P)$ 上の確率変数, $\mathcal{G}$ を $\mathcal{F}$ の部分 σ 加法族, φ を $\mathbf{R}$ 上の凸関数とする. X および $\varphi(X)$ が P 可積分ならば,

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

特に, $p \geq 1$ に対して $X \in L^p(\Omega, P)$ ならば,

$$|E[X|\mathcal{G}]|^p \leq E[|X|^p|\mathcal{G}].$$

証明　φ は凸だから, 任意の $a \in \mathbf{R}$ に対して, ある $c \in \mathbf{R}$ が存在し

$$\varphi(x) \geq \varphi(a) + c(x - a) \quad (x \in \mathbf{R}).$$

更に, ある $\mathcal{G}$ 可測関数 $\tilde{c}$ が存在し

$$\varphi(X) \geq \varphi(E[X|\mathcal{G}]) + \tilde{c}(X - E[X|\mathcal{G}])$$

となることが分かる.
よって, 上の定理の (1)〜(3) より,

$$\begin{aligned} E[\varphi(X)|\mathcal{G}] &\geq E[\varphi(E[X|\mathcal{G}])|\mathcal{G}] + E[\tilde{c}(X - E[X|\mathcal{G}])|\mathcal{G}] \\ &= \varphi(E[X|\mathcal{G}]) + \tilde{c}E[X - E[X|\mathcal{G}]|\mathcal{G}] \\ &= \varphi(E[X|\mathcal{G}]) + \tilde{c}\,(E[X|\mathcal{G}] - E[X|\mathcal{G}]) \\ &= \varphi(E[X|\mathcal{G}]). \end{aligned}$$

□

関連事項 7. 凸関数の片側微分可能性

開区間で定義された凸関数は右導関数および左導関数をもち, 特に連続関数となる.
φ を開区間 I で定義された凸関数とし, $x, y, z \in I$ を

$$x < y < z$$

をみたすように選んでおく. このとき, 凸関数の性質より,

$$\frac{\varphi(y) - \varphi(x)}{y - x} \leq \frac{\varphi(z) - \varphi(x)}{z - x} \leq \frac{\varphi(z) - \varphi(y)}{z - y}$$

がなりたつ. 異なる $x_1, x_2 \in I$ に対して

$$\psi(x_1, x_2) = \frac{\varphi(x_2) - \varphi(x_1)}{x_2 - x_1}$$

とおくと, 上の不等式は

$$\psi(x, y) \leq \psi(x, z) \leq \psi(y, z)$$

と表すことができる. 左側の不等式より, 関数 $\psi(x_1, x_2)$ は x_2 が x_1 よりも大きい値から x_1 へ近づくにつれて単調に減少し, φ の定義域が開区間であることから, その値は下に有界である.
よって, φ は任意の $x \in I$ に対して右微分係数

$$\varphi'_+(x) = \lim_{h \to +0} \frac{\varphi(x + h) - \varphi(x)}{h}$$

をもつ. ここで, $x, y, z, w \in I$ を

$$x < y < z < w$$

をみたすように選んでおくと,

$$\psi(x, y) \leq \psi(x, z) \leq \psi(z, w)$$

がなりたつから, 右導関数 φ'_+ は単調増加である. 更に,

$$\begin{aligned} \varphi'_+(y) &\leq \psi(y, z) \\ &\to \psi(x, z) \quad (y \to x + 0) \\ &\to \varphi'_+(x) \quad (z \to x + 0) \end{aligned}$$

となるから, φ'_+ は右連続である. よって, φ'_+ は Borel 可測関数となる.
同様に, φ は任意の $x \in I$ に対して左微分係数 $\varphi'_-(x)$ をもち, 左導関数 φ'_- は単調増加で左連続な Borel 可測関数となる. また,

$$\varphi'_-(x) \leq \varphi'_+(x)$$

がなりたつ.
更に, $a \in I$ とし, $c \in \mathbf{R}$ が

$$\varphi'_-(a) \leq c \leq \varphi'_+(a)$$

をみたすとすると, 任意の $x \in I$ に対して

$$\varphi(x) \geq \varphi(a) + c(x - a)$$

がなりたつ.