§4. 確率変数と分布

ここでは, 確率論や統計学における基本的用語について, 確率変数とその分布を中心に述べていこう.

まず, $(\Omega, \mathcal{F})$ を可測空間とする. すなわち, Ω は集合, $\mathcal{F}$ は Ω の部分集合の集まりで, 次の (1)〜(3) をみたす.

(1) $\Omega \in \mathcal{F}$.

(2) $A \in \mathcal{F}$ ならば, $A^c \in \mathcal{F}$.

(3) $A_n \in \mathcal{F}$ $(n \in \mathbf{N})$ ならば, $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

このとき, $\mathcal{F}$ を Ω 上の σ 加法族という.

また, P を $(\Omega, \mathcal{F})$ 上の確率測度とする. すなわち, P は $\mathcal{F}$ で定義された関数で, 次の (1), (2) をみたす.

(1) 任意の $A \in \mathcal{F}$ に対して $0 \leq P(A) \leq 1$ で, $P(\Omega) = 1$.

(2) $A_n \in \mathcal{F}$ $(n \in \mathbf{N})$, $A_n \cap A_m = \emptyset$ $(n \neq m)$ ならば,

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n) \ (\sigma \text{ 加法性}).$$

このとき, $(\Omega, \mathcal{F}, P)$ を確率空間という.

定理　$(\Omega, \mathcal{F}, P)$ を確率空間とすると, 次の (1)〜(4) がなりたつ.

(1) $A \in \mathcal{F}$ ならば, $P(A^c) = 1 - P(A)$.

(2) $A_n \in \mathcal{F}$ $(n \in \mathbf{N})$ ならば, $P\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} P(A_n)$ (劣加法性).

(3) $A_n \in \mathcal{F}$, $A_n \subset A_{n+1}$ $(n \in \mathbf{N})$ ならば, $P\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} P(A_n)$.

(4) $A_n \in \mathcal{F}$, $A_n \supset A_{n+1}$ $(n \in \mathbf{N})$ ならば, $P\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} P(A_n)$.

簡単のため, $\mathbf{R}$ に値をとる確率変数について述べよう. $\mathbf{R}$ の Borel 集合族を $\mathcal{B}(\mathbf{R})$ と表す. すなわち, $\mathcal{B}(\mathbf{R})$ は $\mathbf{R}$ のすべての開集合を含む最小の σ 加法族である.

定義　関数

$$X : \Omega \to \mathbf{R}$$

は $\mathcal{F}$ 可測のとき, すなわち任意の $A \in \mathcal{B}(\mathbf{R})$ に対して

$$X^{-1}(A) \in \mathcal{F}$$

がなりたつとき, $(\Omega, \mathcal{F}, P)$ 上の確率変数という.

X を $(\Omega, \mathcal{F}, P)$ 上の確率変数とする. X が P 可積分のとき,

$$E[X] = \int_{\Omega} X(\omega) P(d\omega)$$

とおき, $E[X]$ を X の平均値または期待値という.

X および X^2 が P 可積分のとき,

$$V(X) = \int_\Omega (X(\omega) - E[X])^2 P(d\omega)$$

とおき, $V(X)$ を X の分散という.

確率空間上の確率変数に対して, 分布という $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$ 上の確率測度を定めることができる. X を $(\Omega, \mathcal{F}, P)$ 上の確率変数とし, $A \in \mathcal{B}(\mathbf{R})$ に対して

$$\mu_X(A) = P(X^{-1}(A))$$

とおく.

定理　μ_X は $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$ 上の確率測度.

証明　まず, P は確率測度だから, μ_X の定義より, 任意の $A \in \mathcal{B}(\mathbf{R})$ に対して

$$0 \leq \mu_X(A) \leq 1.$$

また,

$$\begin{aligned}\mu_X(\mathbf{R}) &= P(X^{-1}(\mathbf{R})) \\ &= P(\Omega) \\ &= 1.\end{aligned}$$

次に, $A_n \in \mathcal{B}(\mathbf{R})$ $(n \in \mathbf{N})$ とし,

$$E_n = X^{-1}(A_n)$$

とおくと, X は確率変数だから,

$$E_n \in \mathcal{F}.$$

更に,

$$A_n \cap A_m = \emptyset \quad (n \neq m)$$

とすると,

$$E_n \cap E_m = \emptyset \quad (n \neq m).$$

P は確率測度だから,

$$\begin{aligned}\mu_X\left(\bigcup_{n=1}^\infty A_n\right) &= P\left(X^{-1}\left(\bigcup_{n=1}^\infty A_n\right)\right) \\ &= P\left(\bigcup_{n=1}^\infty E_n\right) \\ &= \sum_{n=1}^\infty P(E_n) \\ &= \sum_{n=1}^\infty \mu_X(A_n).\end{aligned}$$

よって, μ_X は $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$ 上の確率測度. □

上の μ_X を X の分布という.

確率変数の期待値は分布を用いて表すことができる．より一般的な形で述べると次がなりたつ．

定理　X を $(\Omega, \mathcal{F}, P)$ 上の確率変数, φ を Borel 可測関数とする．$\varphi(X)$ が P 可積分ならば, φ は μ_X 可積分で,

$$E[\varphi(X)] = \int_{\mathbf{R}} \varphi(x)\mu_X(dx).$$

証明　まず, 確率変数と Borel 可測関数の合成は確率変数となることに注意しよう．
φ が非負のとき, $n \in \mathbf{N}$ に対して

$$\varphi_n = \sum_{k=0}^{n2^n} \frac{k}{2^n} \chi_{\varphi^{-1}\left(\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right)\right)}$$

とおく．ただし, 集合 A の定義関数を χ_A と表すことにする．
このとき, φ_n は非負単関数の単調増加列で,

$$\lim_{n\to\infty} \varphi_n(x) = \varphi(x) \quad (x \in \mathbf{R}).$$

よって, $\varphi_n(X)$ も非負単調増加の確率変数列で,

$$\lim_{n\to\infty} \varphi_n(X(\omega)) = \varphi(X(\omega)) \quad (\omega \in \Omega).$$

ここで,

$$\begin{aligned}
E(\varphi_n(X)) &= \sum_{k=0}^{n2^n} \frac{k}{2^n} P\left((\varphi \circ X)^{-1}\left(\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right)\right)\right) \\
&= \sum_{k=0}^{n2^n} \frac{k}{2^n} \mu_X\left(\varphi^{-1}\left(\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right)\right)\right) \\
&= \int_{\mathbf{R}} \varphi_n(x)\mu_X(dx).
\end{aligned}$$

したがって, $n \to \infty$ とすると, 積分の定義より,

$$E[\varphi(X)] = \int_{\mathbf{R}} \varphi(x)\mu_X(dx).$$

φ が非負でないときは, φ を 2 つの非負可測関数の差として表して考えればよい．　□

例　X を $(\Omega, \mathcal{F}, P)$ 上の確率変数, μ_X を X の分布とする．
X の期待値が存在すると仮定し, $m = E(X)$ とおくと, 上の定理より,

$$m = \int_{\mathbf{R}} x\mu_X(dx).$$

更に, X^2 の期待値が存在すると仮定すると,

$$V(X) = \int_{\mathbf{R}} (x-m)^2 \mu_X(dx).$$

関連事項 4. 確率変数列の収束

確率変数の列に対しては様々な収束の概念が用いられる.
$(\Omega, \mathcal{F}, P)$ を確率空間, $\{X_n\}_{n=1}^{\infty}$ を Ω 上の確率変数列, X を Ω 上の確率変数とする.
まず,

$$P\left(\left\{\omega \in \Omega \,\middle|\, \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\}\right) = 1$$

がなりたつとき, $\{X_n\}_{n=1}^{\infty}$ は X に概収束するという. このとき,

$$X_n \to X \quad P\text{-a.s.}$$

などと表す.
次に, 任意の $\varepsilon > 0$ に対して

$$\lim_{n\to\infty} P\left(\{\omega \in \Omega \,|\, |X_n(\omega) - X(\omega)| > \varepsilon\}\right) = 0$$

がなりたつとき, $\{X_n\}_{n=1}^{\infty}$ は X に確率収束するという.
$\{X_n\}_{n=1}^{\infty}$ が X に概収束するならば, $\{X_n\}_{n=1}^{\infty}$ は X に確率収束することが分かる. 逆に, $\{X_n\}_{n=1}^{\infty}$ が X に確率収束するならば, X に概収束する部分列が存在することが分かる.
更に, $p \geq 1$ とし, $X_n, X \in L^p(\Omega, P)$ $(n \in \mathbf{N})$ で,

$$\lim_{n\to\infty} E\left[|X_n - X|^p\right] = 0$$

がなりたつとき, $\{X_n\}_{n=1}^{\infty}$ は X に p 次平均収束する, または L^p 収束するという.
$\{X_n\}_{n=1}^{\infty}$ が X に L^p 収束するならば, $\{X_n\}_{n=1}^{\infty}$ は X に確率収束する. 実際, 任意の $\varepsilon > 0$ に対して不等式

$$\varepsilon^p \chi_{\{\omega\in\Omega \,|\, |X_n(\omega) - X(\omega)| > \varepsilon\}} \leq |X_n - X|^p$$

がなりたつことを用いて, 両辺を積分し, $n \to \infty$ とすればよい. 特に, $\{X_n\}_{n=1}^{\infty}$ が X に L^p 収束するならば, X に概収束する部分列が存在する.
また, $\mathbf{R}$ 上の任意の有界連続関数 φ に対して

$$\lim_{n\to\infty} E[\varphi(X_n)] = E[\varphi(X)]$$

がなりたつとき, $\{X_n\}_{n=1}^{\infty}$ は X に法則収束するという.
$\{X_n\}_{n=1}^{\infty}$ が X に確率収束するならば, $\{X_n\}_{n=1}^{\infty}$ は X に法則収束することが分かる.
上の式は X_n の分布 μ_{X_n} および X の分布 μ_X を用いて,

$$\lim_{n\to\infty} \int_{\mathbf{R}} \varphi(x)\mu_{X_n}(dx) = \int_{\mathbf{R}} \varphi(x)\mu_X(dx)$$

と表すことができる. このことを $\{\mu_{X_n}\}_{n=1}^{\infty}$ は μ_X に弱収束するという. 特に, 法則収束について論じる場合は X_n と X が同じ確率空間上で定義されている必要はない.
$\{X_n\}_{n=1}^{\infty}$ が X に法則収束するならば, ある確率空間上の確率変数列 $\{Y_n\}_{n=1}^{\infty}$ および確率変数 Y で, Y_n の分布が μ_{X_n}, Y の分布が μ_X となるものが存在し, $\{Y_n\}_{n=1}^{\infty}$ は Y に概収束することが知られている. これを Skorokhod の表現定理という.