

中心極限定理

ナンブ キトラ

2023 年 6 月 20 日

この文書では中心極限定理の証明を紹介する。

1 基本事項

ここらへんのことは [1] や [2] を参照しても良い。

定義 1. 位相空間 X に対して、 X の開集合系から生成される完全加法族を $\mathfrak{B}(X)$ と書くことにする。この集合族を**ボレル集合族**と呼ぶ。

定義 2. 測度空間 $(X, \mathfrak{M})$ 上の測度 μ が**確率測度**であるとは、 $\mu(X) = 1$ を満たすときにいう。 μ が確率測度であるとき、測度空間 $(X, \mathfrak{M}, \mu)$ を**確率空間**と呼び、 X の元のことを**事象**と呼ぶ。

定義 3. 確率空間 $(X, \mathfrak{M}, \mu)$ から可測空間 $(Y, \mathfrak{N})$ への可測写像可測写像を $(Y, \mathfrak{N})$ -**値確率変数**と呼ぶ。あるいは紛れがないときには単に**確率変数**と呼んだりもする。

定義 4. 確率空間 $(X, \mathfrak{M}, \mu)$ とその上の確率変数 f について $\int_X f \, d\mu$ を $\mathbb{E}(f)$ と書いたりする。特に $\mathbb{E}(f)$ を f の**期待値**と呼んだり、 f の**平均値**と呼んだりする。そして $n \in \mathbb{Z}_{\geq 1}$ に対して、 $\mathbb{E}(f^n)$ を f の n 次モーメントと呼んだりする。 $\mathbb{E}(f^n) < \infty$ であるとき f は L^n 関数とか言ったりするが、これは測度論と同じ言葉遣いである。また、 $\mathbb{V}(f) = \mathbb{E}((f - \mathbb{E}(f))^2)$ を f の**分散**と呼ぶ。簡単な式変形で $\mathbb{V}(f) = \mathbb{E}(f^2) - (\mathbb{E}(f))^2$ となることがわかる。

定義 5. $(X, \mathfrak{M}, \mu)$ を確率空間とし、 $(Y, \mathfrak{N})$ を可測空間とする。このとき可測写像 $\phi : X \rightarrow Y$ に対して写像 $\phi_*\mu : \mathfrak{N} \rightarrow [0, \infty]$ を $\phi_*\mu(A) = \mu(\phi^{-1}(A))$ と定義するとこれは $(Y, \mathfrak{N})$ 上の測度になる。この測度 $\phi_*\mu$ を ϕ による μ の押し出し測度と呼ぶ。

補題 6. $(X, \mathfrak{M}, \mu)$ を確率空間とし, $(Y, \mathfrak{N})$ を可測空間とする. このとき可測写像 $\phi: X \rightarrow Y$ と任意の可測関数 $f: Y \rightarrow \mathbb{R}$ について

$$\int_X f(\phi(x)) d\mu = \int_Y f(y) d(\phi_*\mu)$$

が成り立つ.

定義 7. $\mathbb{R}$ 上の $\mathfrak{B}(\mathbb{R})$ で定義された一次元ルベーグ測度を $\mathcal{L}^1$ と書くことにする.

補題 8. $(X, \mathfrak{M}, \mu)$ を確率空間とする. $\phi: X \rightarrow [0, \infty]$ を $(X, \mathfrak{M})$ 上の可測関数とする. そして新しく $(X, \mathfrak{M})$ 上の測度 ν を

$$\nu(A) = \int_A \phi d\mu$$

と定義する. すると X 上の可測関数 f について

$$\int_X f d\nu = \int_X f\phi d\mu$$

が成り立つ.

証明. 単関数近似を使うとわかる. □

以下の定理の証明は [3] を参考にしても良い. 証明は省略する.

定理 9 (Prokhorov). X を可分完備距離化可能空間とし, $\mathcal{S} \subset \mathcal{P}(X)$ とする. このとき以下は同値である.

- (1) $\mathcal{S}$ は相対コンパクトである.
- (2) $\mathcal{S}$ は緊密である.

以下の二つの補題については [4] を参考にしてもいい. 証明は省く.

補題 10 (ディリクレ積分). 以下が成り立つ.

$$\lim_{R \rightarrow \infty} \int_0^R \frac{\sin x}{x} dx = \frac{\pi}{2}$$

さらに

$$\lim_{R \rightarrow \infty} \int_0^R \frac{\sin(ax)}{x} dx = \begin{cases} \frac{\pi}{2} & \text{if } a > 0 \\ -\frac{\pi}{2} & \text{if } a < 0 \\ 0 & \text{if } a = 0 \end{cases}$$

が成り立つ.

補題 11. $m \in [0, \infty)$ で $\sigma \in (0, \infty)$ とする. $f_{m,\sigma} : \mathbb{R} \rightarrow [0, \infty)$ を $N(m, \sigma)$ の確率密度関数とする. つまり

$$f_{m,\sigma}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$$

である. このとき

$$F_{m,\sigma}(x) = \int_{\mathbb{R}} f_{m,\sigma}(y) \exp(\sqrt{-1}xy) dy$$

とおくことにする. このとき

$$\int_{\mathbb{R}} f_{m,\sigma}(y) \exp(\sqrt{-1}xy) dy = \exp\left(\sqrt{-1}mx - \frac{\sigma^2 x^2}{2}\right)$$

つまり

$$F_{m,\sigma}(x) = \exp\left(\sqrt{-1}mx - \frac{\sigma^2 x^2}{2}\right)$$

となる. 特に

$$F_{0,\sigma} = \exp\left(-\frac{\sigma^2 x^2}{2}\right)$$

であり,

$$F_{0,1} = \exp\left(-\frac{x^2}{2}\right)$$

である.

定義 12. $m \in [0, \infty)$ で $\sigma \in (0, \infty)$ とし,

$$f_{m,\sigma}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$$

とする.

$(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ 上の測度 ν を

$$\nu(A) = \int_A f_{m,\sigma} d\mathcal{L}^1$$

と定義する. この測度 ν を**平均が m で分散が σ の正規分布**と言ったりする. 一般に確率空間 $(X, \mathfrak{M}, \mu)$ 上の確率変数 $f : X \rightarrow \mathbb{R}$ について, 押し出し測度 $f_*\mu$ が平均が m で分散が σ の正規分布に一致するとき f は**平均が m で分散が σ の正規分布に従う**などと言ったりする. また, 平均が m で分散が σ の正規分布を $N(m, \sigma)$ で表し, f は $N(m, \sigma)$ に従うと言ったりもする.

定義 13 (分布関数). $(\mathbb{R}, \mathfrak{B})$ 上の確率測度 μ に対して関数 $F_\mu : \mathbb{R} \rightarrow \mathbb{R}$ を $F_\mu(x) = \mu((-\infty, x])$ と定義する. この F_μ を μ の**分布関数**と呼ぶ一般に確率空間 $(X, \mathfrak{M}, \mu)$ とその上の確率変数 f に対して, $F_{f*\mu}$ を f の**分布関数**と言ったりする. 分布関数のことを**累積分布関数**とも呼ぶ.

定理 14. $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ 上の二つの確率測度 μ, ν に対して $F_\mu = F_\nu$ ならば, $\mu = \nu$ である.

証明. 任意の実数 $a, b \in \mathbb{R}$ が $a < b$ を満たすならば $\mu((a, b]) = F_\mu(b) - F_\mu(a)$ であり, また $\nu((a, b]) = F_\nu(b) - F_\nu(a)$ なので, $F_\mu = F_\nu$ であることから集合族 $\{(a, b] \mid a < b\}$ の上で μ と ν は一致することがわかる. この集合族が生成する完全加法族はボレル集合族に一致するので $\mu = \nu$ である. $\square$

2 特性関数

この節では確率測度の特性関数の定義と基本的な性質を紹介する.

定義 15. μ を $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ 上の測度とする. このとき関数 $\phi_\mu : \mathbb{R} \rightarrow \mathbb{R}$ を

$$\phi_\mu(x) = \int_{\mathbb{R}} \exp(\sqrt{-1}x \cdot y) d\mu(y)$$

と定義する. この関数を μ の**特性関数**と呼ぶ.

定理 16 (Lévy の反転公式). $\mu \in \mathcal{P}(\mathbb{R})$ とする. 記述を簡単にするために $\phi = \phi_\mu$, $F = F_\mu$ とおく, このとき F_μ の任意の連続点 a, b について $a < b$ ならば

$$F_\mu(b) - F_\mu(a) = \lim_{R \rightarrow \infty} \frac{1}{2\pi} \int_{-R}^{+R} \frac{\exp(-\sqrt{-1}xb) - \exp(-\sqrt{-1}xa)}{-\sqrt{-1}x} \phi(x) d\mathcal{L}^1(x)$$

が成り立つ.

証明. まず, Fubini—Toneli の定理から任意の $R > 0$ について

$$\begin{aligned} & \int_{-R}^R \frac{\exp(-\sqrt{-1}xb) - \exp(-\sqrt{-1}xa)}{-\sqrt{-1}x} \phi(x) d\mathcal{L}^1(x) \\ &= \int_{-R}^R \int_{\mathbb{R}} \frac{\exp(\sqrt{-1}x(t-b)) - \exp(\sqrt{-1}x(t-a))}{\sqrt{-1}x} d\mu(t) d\mathcal{L}^1(x) \\ &= \int_{\mathbb{R}} \int_{-R}^R \frac{\exp(\sqrt{-1}x(t-b)) - \exp(\sqrt{-1}x(t-a))}{\sqrt{-1}x} d\mathcal{L}^1(x) d\mu(t) \end{aligned}$$

が成り立つ.

ここで, 関数

$$\frac{\cos(x(t-b)) - \cos(x(t-a))}{x}$$

は任意の t, a, b に対して x に関して奇関数なので (さらに x に関して連続にもなる)

$$\int_{-R}^R \frac{\cos(x(t-b)) - \cos(x(t-a))}{x} d\mathcal{L}^1(x) = 0$$

となる. ゆえに

$$\begin{aligned} & \int_{-R}^R \frac{\exp(\sqrt{-1}x(t-b)) - \exp(\sqrt{-1}x(t-a))}{\sqrt{-1}x} d\mathcal{L}^1(x) \\ &= \int_{-R}^R \frac{\sin(x(t-a)) - \sin(x(t-b))}{x} d\mathcal{L}^1(x) \end{aligned}$$

が成り立つ. ここで,

$$f_R(a) = \int_0^R \frac{\sin(ax)}{x} d\mathcal{L}^1(x)$$

と定義すると

$$\int_{-R}^R \frac{\sin(x(t-a)) - \sin(x(t-b))}{x} d\mathcal{L}^1(x) = 2(f_R(t-a) - f_R(t-b))$$

なので

$$\begin{aligned} & \int_{-R}^R \frac{\exp(-\sqrt{-1}xb) - \exp(-\sqrt{-1}xa)}{-\sqrt{-1}x} \phi(x) d\mathcal{L}^1(x) \\ &= 2 \int_{\mathbb{R}} f_R(t-a) - f_R(t-b) d\mu(t) \end{aligned}$$

である.

ディリクレ積分の式 (補題 10)

$$\lim_{R \rightarrow \infty} f_R(a) = \begin{cases} \frac{\pi}{2} & \text{if } a > 0 \\ -\frac{\pi}{2} & \text{if } a < 0 \\ 0 & \text{if } a = 0 \end{cases}$$

から, 任意の $t, c \in \mathbb{R}$ に関して

$$\lim_{R \rightarrow \infty} f_R(t-c) = \frac{\pi}{2} (\chi_{(c, \infty)} - \chi_{(-\infty, c)})$$

が成り立つ。よって、 $\sup_x |f_R(x)| < \infty$ と優収束定理から

$$\begin{aligned}
& \lim_{R \rightarrow \infty} \int_{-R}^R \frac{\exp(-\sqrt{-1}xb) - \exp(-\sqrt{-1}xa)}{-\sqrt{-1}x} \phi(x) d\mathcal{L}^1(x) \\
&= \lim_{R \rightarrow \infty} 2 \int_{\mathbb{R}} f_R(t-a) - f_R(t-b) d\mu(t) \\
&= \pi \int_{\mathbb{R}} \chi_{(a,\infty)}(t) - \chi_{(-\infty,a)}(t) - \chi_{(b,\infty)}(t) + \chi_{(-\infty,b)}(t) d\mu(t) \\
&= \pi \int_{\mathbb{R}} \chi_{(a,b]}(t) + \chi_{[a,b)}(t) d\mu(t) \\
&= \pi \int_{\mathbb{R}} \chi_{(a,b]}(t) + \chi_{(a,b]}(t) + \chi_{\{a\}}(t) - \chi_{\{b\}}(t) d\mu(t) \\
&= 2\pi \int_{\mathbb{R}} \chi_{(a,b]}(t) d\mu(t) = 2\pi(F(b) - F(a))
\end{aligned}$$

ここで、 p が F の連続点ならば $\int_{\mathbb{R}} \chi_{\{p\}}(t) d\mu(t) = \mu(\{p\}) = 0$ であることを用いた。これで証明が終わる。 $\square$

注意. 定理の主要な等式は以下である。

$$F_\mu(b) - F_\mu(a) = \frac{1}{2\pi} \lim_{R \rightarrow \infty} \int_{-R}^{+R} \frac{\exp(-\sqrt{-1}xb) - \exp(-\sqrt{-1}xa)}{-\sqrt{-1}x} \phi(x) d\mathcal{L}^1(x)$$

この式は

$$F_\mu(b) - F_\mu(a) = \frac{1}{2\pi} \int_a^b \int_{\mathbb{R}} \phi(x) \exp(-\sqrt{-1}x \cdot y) d\mathcal{L}^1(x) d\mathcal{L}^1(y)$$

とみなすことができる。特性関数とは μ の“確率密度関数”のフーリエ変換であり、上の式はその特性関数をフーリエ逆変換してのちに $[a, b]$ 上で積分した量に等しいと考えることができるので、その値は $F(b) - F(a)$ に等しいという観察ができる。しかし一般に特性関数は急減少関数などのフーリエ変換について良い振る舞いをする関数になるとは限らないので、

$$F_\mu(b) - F_\mu(a) = \frac{1}{2\pi} \int_a^b \int_{\mathbb{R}} \phi(x) \exp(-\sqrt{-1}x \cdot y) d\mathcal{L}^1(x) d\mathcal{L}^1(y)$$

が成り立つのかどうか、もしくはそもそも $\phi(x)$ が $\mathbb{R}$ 上で可積分かどうかすらわからないので、上の観察を少し修正して、上の定理の様に主値積分で累積分布関数の情報を引き出す必要があったわけである。

補題 17. 以下の不等式が成り立つ.

$$3 < \pi < 4$$

特に $\pi/4 < 1$ であり, $\sqrt{2} < \pi/2$ である.

証明. 関数 f を $f(t) = (1-t^2)^{-1/2}$ と定義する. $t \in [0, 1/2]$ のとき, $1 \leq f(t) \leq 2/\sqrt{3} \leq 4/3$ である. さらに $1 < f(t) < 4/3$ となる点 $t \in [0, 1/2]$ が存在する. 例えば $t = 1/4$ においてこの不等式が成り立つ. 関数 $f(t)$ は連続だから,

$$\frac{\pi}{6} = \int_0^{1/2} \frac{1}{\sqrt{1-t^2}} dt > \int_0^{1/2} dt = \frac{1}{2}$$

なので, $3 < \pi$ がわかる. また,

$$\frac{\pi}{6} = \int_0^{1/2} \frac{1}{\sqrt{1-t^2}} dt < \int_0^{1/2} \frac{4}{3} dt = \frac{2}{3}$$

なので $\pi < 4$ がわかる. これらの積分の不等式において真の不等式が得られるのは f が連続であり, $1 < f(t) < 4/3$ となる t が存在するからである. $\square$

補題 18. 実数 a が $|a| \geq 1$ を満たすならば,

$$|\sin a| \leq |a| \sin(1)$$

が成り立つ.

証明. $0 \leq a$ のときだけ証明すれば良い. $f(x) = x \sin(1) - \sin x$ とおく. $x \in [\pi/4, \pi/2]$ のとき $f'(x) = \sin(1) - \cos x$ の符号を調べる. $\pi/4 < 1$ なので, $1/\sqrt{2} < \sin 1$ であり, $[\pi/4, \pi/2]$ 上では $\cos x$ は単調減少なので $f'(x) = \sin(1) - \cos x > 0$ である. そして $f(1) = 0$ なので, $x \in [1, \pi/2]$ のとき $0 < f(x)$ である. $\frac{1}{\sin(1)} < \sqrt{2} < \frac{\pi}{2}$ なので, 特に $\pi/2 < x$ ならば, $x \sin 1 > 1$ であるから, この場合も $0 < f(x)$ が成り立つ. よって $1 \leq x$ ならば $\sin x \leq x \sin(1)$ である. $\square$

命題 19. 任意の $\mu, \nu \in \mathcal{P}(\mathbb{R})$ について $\phi_\mu = \phi_\nu$ ならば $\mu = \nu$ である.

証明. 定理 16 より $\phi_\mu = \phi_\nu$ ならば, 可算個の点を除いた全ての $x, a \in \mathbb{R}$ について $F_\mu(x) - F_\mu(a) = F_\nu(x) - F_\nu(a)$ である. $a \rightarrow -\infty$ とすると F_μ も F_ν も 0 に限りなく近づくので可算個の点を除いた全ての $x \in \mathbb{R}$ で $F_\mu(x) = F_\nu(x)$ である. 分布関数は右連続な単調関数なので, 結局すべての点 $x \in \mathbb{R}$ で $F_\mu(x) = F_\nu(x)$ となる. ゆえに定理 14 より $\mu = \nu$ となる. $\square$

命題 20. 点列 $\{\mu_n\}_{n \in \mathbb{Z}_{\geq 0}} \subset \mathcal{P}(\mathbb{R})$ と $\mu \in \mathcal{P}(\mathbb{R})$ について $\mu_n \Longrightarrow \mu$ と $\phi_{\mu_n} \rightarrow \phi_\mu$ (各点収束) は同値である.

証明. 弱収束 $\mu_n \Longrightarrow \mu$ が成り立つならば $\phi_{\mu_n} \rightarrow \phi_\mu$ (各点収束) となることは定義からすぐにわかる. 逆を示そう. 記述を簡単にするために $\phi = \phi_\mu$, $\phi_n = \phi_{\mu_n}$ とおくことにする. まず $\{\mu_n\}$ の緊密性を示そう. $M \in (0, \infty)$ とする. 今から

$$\mu_n(\{x \in \mathbb{R} \mid |x| \geq M\})$$

を評価する. 任意の $\epsilon \in (0, \infty)$ と 十分大きい M と 十分大きい $N \in \mathbb{Z}_{\geq 0}$ について $n > N$ ならば $\mu_n(\{x \in \mathbb{R} \mid |x| \geq M\}) \leq \epsilon$ を満たすことを示せば $\{\mu_n\}_{n \in \mathbb{Z}_{\geq 0}}$ が緊密であることがわかる. $|x| \geq M$ について $|x|/M \geq 1$ なので補題から, $\sin(|x|/M) \leq |x|/M \sin(1)$ であるから,

$$1 \leq \frac{1}{1 - \frac{1}{\sin(1)}} \left(1 - \frac{\sin(x/M)}{x/M} \right)$$

が成り立つ. $c = (1 - 1/\sin(1))^{-1}$ と置くと,

$$1 \leq c \left(1 - \frac{\sin(x/M)}{x/M} \right)$$

となる. よって,

$$\chi_{\{|x| \geq M\}} \leq c \left(1 - \frac{\sin(x/M)}{x/M} \right)$$

なので, 任意の $n \in \mathbb{N}$ について

$$\mu_n(\{x \in \mathbb{R} \mid |x| \geq M\}) \leq c \int_{\mathbb{R}} \left(1 - \frac{\sin(x/M)}{x/M} \right) d\mu_n$$

を得る. さて, $x \in \mathbb{R}$ のとき,

$$\frac{1}{2} \int_{-1}^1 \exp(\sqrt{-1}xy) d\mathcal{L}^1(y) = \left[\frac{1}{2\sqrt{-1}x} \exp(\sqrt{-1}xy) \right]_{y=-1}^{y=1} = \frac{\sin x}{x}$$

が成り立つので,

$$\begin{aligned} \int_{\mathbb{R}} \left(1 - \frac{\sin(x/M)}{x/M} \right) d\mu_n(x) &= \int_{\mathbb{R}} \left(1 - \frac{1}{2} \int_{-1}^1 \exp\left(\sqrt{-1} \frac{x}{M} y\right) d\mathcal{L}^1(y) \right) d\mu_n(x) \\ &= \mu_n(\mathbb{R}) - \frac{1}{2} \int_{-1}^1 \phi_n\left(\frac{y}{M}\right) d\mathcal{L}^1(y) \\ &= 1 - \frac{M}{2} \int_{-1/M}^{1/M} \phi_n(y) d\mathcal{L}^1(y) \end{aligned}$$

がわかる．ここで

$$I(M) = 1 - \frac{M}{2} \int_{-1/M}^{1/M} \phi(y) dy$$

$$J_n(M) = \frac{M}{2} \int_{-1/M}^{1/M} \phi(y) - \phi_n(y) dy$$

とおけば,

$$\mu_n(\{x \in \mathbb{R} \mid |x| \geq M\}) \leq c \cdot I(M) + c \cdot J_n(M)$$

である．さて, ϕ は $\phi(0) = 1$ であり連続だから, 積分の平均値の定理から $I(M) \rightarrow 0$ ($M \rightarrow \infty$) となる．つまり, 任意の ε に対して, δ が存在して, $|x| < \delta$ ならば $|\phi(x) - 1| < \varepsilon/2c$ となる．このことから M を $\delta^{-1} < M$ を満たすように大きくとれば,

$$|I(M)| \leq \varepsilon/2$$

が成り立つ．

次に $J_n(M)$ を評価する． M を固定して考える． ϕ_n は ϕ に各点収束している．また, 各 μ_n が確率測度であることと, 特性関数の定義から $[-1/M, 1/M]$ 上では $\{\phi_n\}$ は一様有界である (特に n によらず定数 1 で上から抑えられる)．有界収束定理から,

$$\int_{-1/M}^{1/M} \phi_n(y) dy \rightarrow \int_{-1/M}^{1/M} \phi(y) dy$$

が成り立つので, N を十分大きく取れば $n > N$ のとき $J_n(M) < \varepsilon/2c$ となる．以上から十分大きい M と $N \in \mathbb{N}$ を取れば, $n > N$ となる n について $\mu_n(\mathbb{R} \setminus [-M, M]) \leq \varepsilon$ となるから, 列 $\{\mu_n\}_{n \in \mathbb{N}}$ は緊密になる．Prokhorov の定理 (定理 9) から, $\{\mu_n\}_{n \in \mathbb{N}}$ は弱収束の位相について相対コンパクトである．さて, $\{\mu_n\}_{n \in \mathbb{N}}$ の任意の部分列 $\{\mu_{n_k}\}$ をとると, これもやはり相対コンパクトなので, 収束する部分列が取れる．その部分列の収束先を $\nu \in \mathcal{P}(X)$ とすると, 仮定から $\phi_\nu = \phi_\mu$ がわかる．命題 19 から, $\nu = \mu$ が結論されるので, 結局 $\{\mu_n\}_{n \in \mathbb{N}}$ の任意の部分列は収束する部分列が存在し, その収束先は全て μ である．よって元の列 $\{\mu_n\}_{n \in \mathbb{N}}$ は μ に収束する． $\square$

応用として正規分布の再生性などを証明する．

定理 21. $(X, \mathfrak{M}, \mu)$ を確率空間とし, f をその上の確率変数で $N(m, \sigma^2)$ に従うとする．このとき $af + b$ は $N(am + b, a^2\sigma^2)$ に従う．

証明. $g = af + b$ とおく. すると

$$\begin{aligned}
\phi_g(x) &= \int_{\mathbb{R}} \exp(\sqrt{-1}x \cdot y) \, dg_*\mu(y) = \int_{\mathbb{R}} \exp(\sqrt{-1}x \cdot g(t)) \, d\mu(t) \\
&= \int_{\mathbb{R}} \exp(\sqrt{-1}x \cdot (af(t) + b)) \, d\mu(t) \\
&= \exp(\sqrt{-1}xb) \int_{\mathbb{R}} \exp(\sqrt{-1}ax \cdot f(t)) \, d\mu(t) \\
&= \exp(\sqrt{-1}xb) \int_{\mathbb{R}} \exp(\sqrt{-1}(ax) \cdot y) \, df_*\mu(y) \\
&= \exp(\sqrt{-1}xb) \int_{\mathbb{R}} \exp(\sqrt{-1}(ax) \cdot y) \, df_*\mu(y) \\
&= \exp(\sqrt{-1}xb) \cdot \exp\left(\sqrt{-1}amx - \frac{a^2\sigma^2x^2}{2}\right) \\
&= \exp\left(\sqrt{-1}(am + b)x - \frac{a^2\sigma^2x^2}{2}\right)
\end{aligned}$$

となる. これは $N(am + b, a^2\sigma^2)$ の特性関数なので, 定理が成り立つことがわかる. $\square$

定理 22. $(X, \mathfrak{M}, \mu)$ を確率空間とし, f, g をその上の独立な確率変数でそれぞれ $N(m_1, \sigma_1^2)$ と $N(m_2, \sigma_2^2)$ に従うとする. このとき $f + g$ は $N(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$ に従う.

証明. $h = f + g$ とおくすると

$$\begin{aligned}
\phi_g(x) &= \int_{\mathbb{R}} \exp(\sqrt{-1}x \cdot y) \, dh_*\mu(y) = \int_{\mathbb{R}} \exp(\sqrt{-1}x \cdot h(t)) \, d\mu(t) \\
&= \int_{\mathbb{R}} \exp(\sqrt{-1}x \cdot (f(t) + g(t))) \, d\mu(t) \\
&= \int_{\mathbb{R}} \exp(\sqrt{-1}x \cdot f(t)) \cdot \exp(\sqrt{-1}x \cdot g(t)) \, d\mu(t) \\
&= \int_{\mathbb{R}} \exp(\sqrt{-1}x \cdot f(t)) \, d\mu(t) \cdot \int_{\mathbb{R}} \exp(\sqrt{-1}x \cdot g(t)) \, d\mu(t) \\
&= \phi_f(x) \cdot \phi_g(x) = \exp\left(\sqrt{-1}m_1x - \frac{\sigma_1^2x^2}{2}\right) \cdot \exp\left(\sqrt{-1}m_2x - \frac{\sigma_2^2x^2}{2}\right) \\
&= \exp\left(\sqrt{-1}(m_1 + m_2)x - \frac{(\sigma_1^2 + \sigma_2^2)x^2}{2}\right)
\end{aligned}$$

これは $N(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$ の特性関数なので, 定理が従う. $\square$

3 中心極限定理

このセクションでは中心極限定理を証明する.

まずは次の初等的な補題を示そう.

補題 23. 0 でない $a \in \mathbb{C}$ に収束する数列 $\{a_n\} \subset \mathbb{C}$ と 0 に収束する数列 $\{b_n\} \subset \mathbb{C}$ について, $\lim_{n \rightarrow \infty} a_n^n$ が存在するならば,

$$\lim_{n \rightarrow \infty} \left(a_n + \frac{b_n}{n} \right)^n = \lim_{n \rightarrow \infty} a_n^n$$

が成り立つ.

証明. $a \neq 0$ なので a_n は十分大きい番号で $a_n \neq 0$ である. これを使って

$$\left(a_n + \frac{b_n}{n} \right)^n = a_n^n \left(1 + \frac{b_n/a_n}{n} \right)^n$$

が成り立つので a_n が恒等的に 1 の場合に補題を証明すれば良い. 二項定理により

$$\left(1 + \frac{b_n}{n} \right)^n - 1 = \sum_{k=1}^n \binom{n}{k} \frac{b_n^k}{n^k}$$

である. ところで, 各 k について

$$\binom{n}{k} \frac{|b_n|^k}{n^k} \leq \frac{n \cdot (n-1) \cdots (n-k+1)}{1 \cdot 2 \cdots k} \frac{|b_n|^k}{n^k} \leq \frac{|b_n|^k}{k!}$$

である. ここで, $\exp(x) - 1$ と $|b_n|$, 0 に対してテイラーの定理を用いると, $c_n \in (0, |b_n|)$ が存在して

$$\exp(|b_n|) - 1 = \sum_{k=1}^n \frac{|b_n|^k}{k!} + \frac{|b_n|^{n+1}}{(n+1)!} \exp(c_n)$$

であるから,

$$\left| \left(1 + \frac{b_n}{n} \right)^n - 1 \right| \leq \sum_{k=1}^n \binom{n}{k} \frac{|b_n|^k}{n^k} \leq \sum_{k=1}^n \frac{|b_n|^k}{k!} \leq -(\exp(|b_n|) - 1) + \frac{|b_n|^{n+1}}{(n+1)!} \exp(c_n)$$

である. ここで, $n \rightarrow \infty$ のとき $|b_n| \rightarrow 0$, $\exp(|b_n|) \rightarrow 1$ と $1/(n!) \rightarrow 0$ なので,

$$-(\exp(|b_n|) - 1) + \frac{|b_n|^{n+1}}{(n+1)!} \exp(c_n) \rightarrow 0$$

がわかる. このことから補題が示された. □

中心極限定理は以下の定理である.

定理 24. $(X, \mathfrak{M}_X, \mu)$ を確率空間とする. そして $\{f_n\}_{n \in \mathbb{N}}$ をこの確率空間上の独立な可測関数で, 同分布に従い, $m = \mathbb{E}(f_1) < \infty$ $v = \mathbb{V}(f_1) < \infty$ とする. このとき,

$$Z_n = \frac{\sum_{i=1}^n (f_i - m)}{\sqrt{vn}}$$

で定義される可測関数列 $\{Z_n\}_{n \in \mathbb{N}}$ は $N(0, 1)$ に法則収束する.

証明. 最初から $m = 0$ で $v = 1$ としても良い. ϕ_n を $(f_n)_* \mu$ の特性関数とする. このとき f_i 達は同分布に従うので $\phi_n = \phi_1$ である. この関数を単に ϕ で表すことにする. $v < \infty$ なので各 f_i は L^2 であるから, 優収束定理から ϕ は C^2 級関数であり,

$$\phi(0) = 1, \quad \frac{d}{dx} \phi(0) = 0, \quad \frac{d^2}{dx^2} \phi(0) = -1$$

が満たされる. よってテイラーの定理から任意の $x \in \mathbb{R}$ について, $|t(x)| \leq |x|$ となる $t(x)$ が存在して

$$\phi(x) = 1 + \frac{1}{2} x^2 \phi''(t(x))$$

を満たす. $(Z_n)_* \mu$ の特性関数を ϕ_{Z_n} で表すことにする. すると

$$\begin{aligned} \phi_{Z_n}(x) &= \int_{\mathbb{R}} \exp(\sqrt{-1}x \cdot y) d((Z_n)_* \mu)(y) = \int_X \exp(\sqrt{-1}x \cdot Z_n(t)) d\mu(t) \\ &= \int_X \exp\left(\sqrt{-1}x \cdot \frac{\sum_{i=1}^n f_i(t)}{\sqrt{n}}\right) d\mu(t) \\ &= \int_X \prod_{i=1}^n \exp\left(\sqrt{-1}x \frac{f_i(t)}{\sqrt{n}}\right) d\mu(t) \end{aligned}$$

となる. f_i 達の独立性から各 x について $\exp(\sqrt{-1}x f_i(t))$ 達も独立になるので

$$\begin{aligned} \int_X \prod_{i=1}^n \exp\left(\sqrt{-1}x \frac{f_i(t)}{\sqrt{n}}\right) d\mu(t) &= \prod_{i=1}^n \int_X \exp\left(\sqrt{-1} \frac{x}{\sqrt{n}} f_i(t)\right) d\mu(t) \\ &= \prod_{i=1}^n \int_{\mathbb{R}} \exp\left(\sqrt{-1} \frac{x}{\sqrt{n}} y\right) d((f_i)_* \mu)(y) \\ &= \left(\phi\left(\frac{x}{\sqrt{n}}\right)\right)^n \end{aligned}$$

となる. まとめると

$$\phi_{Z_n}(x) = \left(\phi\left(\frac{x}{\sqrt{n}}\right)\right)^n$$

である．ここで $t(x/\sqrt{n}) = t(x, n)$ と書くことにすると，

$$\phi_{Z_n}(x) = \left(1 + \frac{1}{2n}x^2 \cdot \phi''(t(x, n))\right)^n = \left(1 - \frac{1}{2}x^2 + \frac{1}{2n}x^2(\phi''(t(x, n)) + 1)\right)^n$$

と表すことができる．点 $x \in \mathbb{R}$ を固定する．このとき $|t(x, n)| \leq x/\sqrt{n}$ であるから $\lim_{n \rightarrow \infty} t(x, n) = 0$ となる．すると $a_n = 1 - \frac{1}{2n}x^2$, $b_n = \frac{1}{2}x^2(\phi''(t(x, n)) + 1)$ と考えて先の補題 23 を適応できて，

$$\lim_{n \rightarrow \infty} \left(1 - \frac{1}{2n}x^2 + \frac{1}{2n}x^2(\phi''(t(x, n)) + 1)\right)^n = \lim_{n \rightarrow \infty} \left(1 - \frac{1}{n} \frac{x^2}{2}\right)^n = \exp\left(-\frac{1}{2}x^2\right)$$

を得る． $\exp(-\frac{1}{2}x^2)$ は正規分布 $N(0, 1)$ の特性関数であるからこの式は ϕ_{Z_n} は正規分布 $N(0, 1)$ の特性関数に各点収束することを意味する．よって， Z_n は $N(0, 1)$ に法則収束する． $\square$

4 中心極限定理の別証明

この節では中心極限定理の特性関数を使わない別証明を紹介する．

4.1 与えられた確率変数列と同じ分布を持つ独立確率変数の存在と構成

確率論においては，確率空間そのものよりも確率変数を主体と考えて確率空間を明記せず確率変数を扱うことが多い．これは確率変数を考えるにおいてそれが定義されている確率空間の確率測度ではなくそれを確率変数によって押し出しをした $\mathbb{R}$ 上の測度，つまり確率変数の分布を主に扱うためであろう．確率変数の定義されている確率空間の測度が何であろうと，結局 $\mathbb{R}$ に押し出して考えるので，確率変数と $\mathbb{R}$ 上の確率測度を集中して扱えば良いというわけである．こうした考え方はある種の利点であるとも考えられるのだが，普段は確率論や統計学以外の数学を扱っている人物に関してはこうした取り扱いというのは，命題の中で具体的に何を扱っているのかという事象の把握においてそれなりの不安感を与えるものになるだろう．例えば，中心極限定理の仮定に出てくる同分布独立な確率変数などがそうである．しかしそうした不安というものはやろうと思えば数学的にきちんと処理できるものである（なのであんまり気にしなくて良いという考え方もある）．この節ではまず最初に同分布独立な確率変数の存在，もしくは構成を数学的な意味で厳密にとり行う．

定理 25. $(X, \mathfrak{M}, \mu)$ を確率空間とする. I, J を可算集合かもしくは有限集合とする. $\{f_i\}_{i \in I}$ を $(X, \mathfrak{M}, \mu)$ 上の独立な確率変数とする. α を $\mathbb{R}$ 上の確率ボレル測度とする. このとき確率空間 $(Y, \mathfrak{N}, \nu)$ とこの上の二つの確率変数の列 $\{\phi_i\}_{i \in I}$ と $\{\psi_j\}_{j \in J}$, そして可測写像 $\pi: X \rightarrow Y$ が存在して以下を満たす.

- (1) $\{\phi_i\}_{i \in I}$ は独立である.
- (2) $\{\psi_j\}_{j \in J}$ は独立である.
- (3) $\{\phi_i, \psi_j\}_{i \in I, j \in J}$ は独立である.
- (4) $\pi_* \mu = \nu$
- (5) 任意の $i \in I$ について $\phi_i \circ \pi = f_i$ が成り立つ.
- (6) 任意の $i \in I$ について $(f_i)_*(\mu) = (\phi_i)_*(\nu)$ が成り立つ.
- (7) 任意の $j \in J$ について $(\psi_j)_*(\nu) = \alpha$ が成り立つ.

証明. $Y = X \times \prod_{j \in J} \mathbb{R}$ とおき, $\mathfrak{N} = \mathfrak{M} \otimes \bigotimes_{j \in J} \mathfrak{B}(\mathbb{R})$ として $(Y, \mathfrak{N})$ 上の測度を

$$\nu = \mu \otimes \bigotimes_{j \in J} \alpha$$

と定義する.

$\phi_i: Y \rightarrow \mathbb{R}$ を $(x, (r_j)_{j \in J}) \in Y = X \times \prod_{j \in J} \mathbb{R}$ に対して $\phi_i((x, (r_j)_{j \in J})) = f_i(x)$ と定義する.

$\psi_j: Y \rightarrow \mathbb{R}$ を Y から第 j 成分への射影とする. つまり $\psi_j(x, (r_k)_{k \in J}) = r_j$ とする.

π を Y から X への射影とする.

まず (1) を証明しよう. ϕ_i の定義から任意の $A \in \mathfrak{B}(\mathbb{R})$ について $\phi_i^{-1}(A) = f_i^{-1}(A) \times \prod_{j \in J} \mathbb{R}$ なので, f_i 達の独立性から $\{\phi_i\}_{i \in I}$ の独立性がわかる.

(2) を示そう. $A \in \mathfrak{B}(\mathbb{R})$ に対して $\psi_j^{-1}(A) = X \times A \times \prod_{k \in J, k \neq j} \mathbb{R}$ なので, 直積測度の定義から独立性がわかる.

(3) を示そう. $\phi_i^{-1}(A) = f_i^{-1}(A) \times \prod_{j \in J} \mathbb{R}$ と $\psi_j^{-1}(A) = X \times A \times \prod_{k \in J, k \neq j} \mathbb{R}$, そして直積測度の定義からわかる

(4) は直積測度の定義からわかる.

(5) は ϕ_i の定義からわかる.

(6) は $\phi_i^{-1}(A) = f_i^{-1}(A) \times \prod_{j \in J} \mathbb{R}$ からわかる.

(7) は ψ_j が射影であることと, 直積測度の定義からわかる.

□

上の定理で特に $I = \emptyset$ とした場合に次の系が得られる.

系 26. α を $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ 上の確率測度とする. J を可算集合か有限集合とする. このとき確率空間 $(X, \mathfrak{M}, \mu)$ とこの上の確率変数の列 $\{h_j\}_{j \in J}$ が存在して, この列は独立同分布でその分布は α に等しい.

4.2 中心極限定理の別証明

補題 27. $\{(X_n, \mathfrak{M}_n, \mu_n)\}_{n \in \mathbb{Z}_{\geq 0}}$ を確率空間として, ϕ_n は $(X_n, \mathfrak{M}_n)$ 上の確率変数とする. ψ は $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ 上の確率変数とする. S を $(C_b(\mathbb{R}, \mathbb{R}), \|\cdot\|_\infty)$ の稠密部分集合とする. を満たす. このとき任意の $f \in S$ に対して

$$\lim_{n \rightarrow \infty} \mathbb{E}(f(\phi_n)) = \mathbb{E}(f(\psi))$$

が成り立つならば ϕ_n は ψ に法則収束する.

証明. $\epsilon \in (0, \infty)$ を任意に与える. $\nu = \psi_* \mathcal{L}^1$ とおく. 任意に $f \in C_b(\mathbb{R}, \mathbb{R})$ を与える. 仮定から $\|f - g\| < \epsilon$ となる $g \in S$ が存在する. このことから任意の n について $|\mathbb{E}(f(\phi_n)) - \mathbb{E}(g(\phi_n))| \leq \epsilon$ と $|\mathbb{E}(f(\psi)) - \mathbb{E}(g(\psi))| \leq \epsilon$ が成り立つ. また仮定より, 十分大きい n について $|\mathbb{E}(g(\phi_n)) - \mathbb{E}(g(\psi))| < \epsilon$ となる. よって

$$|\mathbb{E}(f(\phi_n)) - \mathbb{E}(f(\psi))| \leq 3\epsilon$$

となる. 以上から ϕ_n は ψ に法則収束することがわかる. □

定理 28. $(X, \mathfrak{M}, \mu)$ を確率空間とし, f_n をその上の独立同分布な確率変数の列で平均値が 0, 分散が 1 とする. このとき新たに確率変数の列 θ_n を

$$\theta_n = \frac{f_1 + \cdots + f_n}{\sqrt{n}}$$

と定義する. すると θ_n は正規分布に法則収束する.

証明. 補題を用いて ϕ_n と, 正規分布に従う確率変数の列 ψ_n をとる. そして新たに確率変数を

$$\zeta_{n,i} = \frac{\phi_1 + \cdots + \phi_{i-1} + \psi_{i+1} + \cdots + \psi_n}{\sqrt{n}}$$

$$\eta_{n,i} = \frac{\phi_1 + \cdots + \phi_i + \psi_{i+1} + \cdots + \psi_n}{\sqrt{n}}$$

と定義する. ここで $\zeta_{n,i}$ の定義に ϕ_i が含まれていないのは誤植ではなく, そういう定義である. 後程 $\zeta_{n,i}$ と ϕ_i が独立であることを使う. さて正規分布の再生性から $\eta_{n,0}$ が正

規分布に従うことに注意しよう．証明の方針としては $\eta_{n,n}$ が $\eta_{n,0}$ に法則収束することを示すのだが，そのために $\eta_{n,i}$ が $\eta_{n,i-1}$ に法則収束することを $\zeta_{n,i}$ をうまく利用して証明する．これを n 段階繰り返せば中心極限定理が証明できるわけである． S を $C_b(\mathbb{R}, \mathbb{R})$ の C^∞ 級関数全体とする． S は $C_b(\mathbb{R}, \mathbb{R})$ の中で稠密になっていることに注意しよう（軟化子を使う）．任意の $f \in S$ について $\mathbb{E}(f(\eta_{n,n}))$ が $\mathbb{E}(f(\eta_{n,0}))$ に収束することを示せばよい．まず確率変数を

$$\kappa_{n,i} = f(\eta_{n,i}) - f(\zeta_{n,i}) - \frac{f^{(1)}(\zeta_{n,i}) \cdot \phi_i}{\sqrt{n}} - \frac{f^{(2)}(\zeta_{n,i}) \cdot \phi_i^2}{2n}$$

と定義する．二階微分に関するテイラーの定理からある $\tau_{n,i}$ が存在して

$$\kappa_{n,i} = \frac{(f^{(2)}(\tau_{n,i}) - f^{(2)}(\zeta_{n,i})) \cdot \phi_i^2}{2n}$$

が成り立つ．この式は実際は各 $y \in Y$ に対して $\kappa_{n,i}(y)$ と書くべきだが，省略して書いている．任意に $\epsilon \in (0, \infty)$ を与える．すると $f^{(3)}$ は有界なのである δ が存在して $\delta \cdot \sup_{t \in \mathbb{R}} |f^{(3)}(t)| < \epsilon$ となる．すると $|x - y| < \delta$ ならば $|f^{(2)}(x) - f^{(2)}(y)| < \epsilon$ となる． $A = \{x \in Y \mid |\phi_i(y)| \leq \delta\sqrt{n}\}$ とする．このとき

$$\mathbb{E}(|\kappa_{n,i}| \cdot \chi_A) \leq \frac{\epsilon}{2n} \mathbb{E}(\phi_i^2) = \frac{\epsilon}{2n}$$

となる．また， $M = \sup |f^{(2)}|$ とおくと

$$\mathbb{E}(|\kappa_{n,i}| \cdot \chi_{Y \setminus A}) \leq \frac{M}{n} \mathbb{E}(|\phi_i| \cdot \chi_{Y \setminus A})$$

となる． A は n に依存しているが， $n \rightarrow \infty$ とすると $\chi_{Y \setminus A} \rightarrow 0$ となる．十分大きい n を取れば $\mathbb{E}(|\phi_i| \cdot \chi_{Y \setminus A}) \leq \epsilon/2(M+1)$ となる．また

$$f(\eta_{n,i}) - f(\zeta_{n,i}) - \frac{f^{(2)}(\zeta_{n,i}) \cdot \phi_i^2}{2n} = \kappa_{n,i} + \frac{f^{(1)}(\zeta_{n,i}) \cdot \phi_i}{\sqrt{n}}$$

なので，

$$\begin{aligned} \left| \mathbb{E}(f(\eta_{n,i})) - \mathbb{E}(f(\zeta_{n,i})) - \mathbb{E}\left(\frac{f^{(2)}(\zeta_{n,i}) \cdot \phi_i^2}{2n}\right) \right| &\leq |\mathbb{E}(\kappa_{n,i})| + \left| \mathbb{E}\left(\frac{f^{(1)}(\zeta_{n,i}) \cdot \phi_i}{\sqrt{n}}\right) \right| \\ &\leq \frac{\epsilon}{n} + \left| \mathbb{E}\left(\frac{f^{(1)}(\zeta_{n,i})}{\sqrt{n}}\right) \right| \cdot \mathbb{E}(\phi_i) \\ &= \frac{\epsilon}{n} \end{aligned}$$

を得る．同じ論法が $\eta_{n,i-1}$ と $\zeta_{n,i}$ にも適応できるので

$$\left| \mathbb{E}(f(\eta_{n,i-1})) - \mathbb{E}(f(\zeta_{n,i})) - \mathbb{E}\left(\frac{f^{(2)}(\zeta_{n,i}) \cdot \phi_i^2}{2n}\right) \right| \leq \frac{\epsilon}{n}$$

となる．よって、

$$|\mathbb{E}(f(\eta_{n,i-1})) - \mathbb{E}(f(\eta_{n,i}))| \leq \frac{2\epsilon}{n}$$

を得る．ゆえに

$$|\mathbb{E}(f(\eta_{n,0})) - \mathbb{E}(f(\eta_{n,n}))| \leq 2\epsilon$$

となるので、中心極限定理が示された。 □

5 注釈

5.1 動的概念と静的概念

まず結論から話すと、理由は知らないが、まあとにかく数学は静的概念しか扱えないということである。言いたいことがあんまり纏まってないので、確率変数が出てくるところから読むといいと思います。

ここからの文章で用いる「動的概念」とは何かしらの一連の動きや働きそのものを指し示す概念付けのことを指す。例えば物体の運動というものも運動という一連の動きを指し示すものとして概念付けされている。一方で「静的概念」とは動的概念ではないもののことである。我々の世界は時間変化するので、この静的概念の具体例を挙げるのはとても難しいのだが、 $1+1=2$ は静的概念であるし、何かしらの留保はつくかも知れないが、りんごは赤色であるという言及もまあ多分静的概念である。このように例示をすると、そもそも静的概念というのは観念の中にしかないものであるから、一番最初に言った結論が成り立つのは当然のことなのかも知れない。どちらかという、静的概念がどうなっているのかということよりは、動的概念がどうなっているのかということがこれからの作用点になる。動的概念は現実世界だけではなく、我々の頭の中、観念の中にも現れる。ここからやっと数学の話になります。高校数学においては極限の定義は以下のように行う。つまり、数列 $a_0, a_1, \dots, a_n, \dots$ が a に収束するとは番号 i を大きくして行ったときに a_i がどんどん a に近づくときにいう。他方で大学などでは以下のように定義する。つまり数列 $\{a_i\}_{i \in \mathbb{Z}_{\geq 0}}$ が a に収束するとは、任意の $\epsilon > 0$ に対して $N \in \mathbb{Z}_{\geq 0}$ が存在して任意の $n \in \mathbb{Z}_{\geq 0}$ が $n \geq N$ を満たすならば $|a_n - a| < \epsilon$ を満たすという定義になる。この二つ

の概念はそれぞれ動的概念と静的概念になっている。前者はどんどん近づくという定義付けをしている。なぜかは知らないがこの定義の仕方だと厳密に扱えないみたいなので、大学では後者のように静的に定義する。後者においてはどんどん近づくという観念は身を潜め、何か ϵ を与えたら十分大きい番号全てで $|a_n - a| < \epsilon$ という言い方になっている。どんどん近づくという言い方をせずに十分大きい番号**すべて**で a_n と a の差が与えられた ϵ の範囲に収まるという言い換えをすることによって動的概念を排除している。さて一番最初に話に戻るが、この収束の話において、「収束の概念を**厳密**に扱うためにはイプシロンデルタ論法」が必要という言い方がされるが、これは確かに正しいのだけど、「厳密に扱うこと」、「数学的に扱うこと」、「動的概念を排除すること」、これらが全て同じことを指しているように見える。そしてその理由は不明である。とにかく、何かしらのことを数学的に扱おうとすると、動的概念を排除するように定義をすとなんだかうまくいくと思う。そうすとなぜか数学的に扱えるようになる。逆に言えば数学は静的概念しか取り扱えないということでもある。(しかしこれは「数学は数学的概念しか扱えない」ということと同じでないだろうか)。

それはさておき、この文書は確率論を扱うためのものなので、確率論に現れる動的概念と静的概念の対立が顕著なものを説明しよう。それは**確率変数**である。世の中には測度論的ではない確率論もあるようで、測度論的確率論が確率論の全てではないのだけど、この文章では数学的な確率論とは測度論的確率論ということになっているので、以下ではそのつもりで記述している。数学的に厳密ではない(**と言われている**) 確率変数の取り扱いはおそらく以下のようなものになるのではないのだろうか？確率変数 X とは、刻一刻と値をさまざまに**ランダム**に取り続ける変数のことである。つまり X は放っておくと $X = 1$ だったり、 $X = 10!$ とか勝手に値がランダムに変わる変数のことである。これは動的概念である。ちなみに静的には(数学的には) 確率変数とはただ単に確率空間 (Ω, μ) 上の(大抵の場合は) 実数値を取る可測関数のことである。確率変数を可測関数と思ってない人は上で述べたように確率変数を認知しているのだと思うし、おそらくこれは人類がプリミティブに認知した確率変数の定義である。この説明において確率変数の分布とは X の取る値を観察したときの値の分布である。つまり X は刻一刻と値を変化させるけれども、 $X = a$ となる瞬間を記録して十分長い時間観察すれば $X = a$ となる確率がわかるということである。可測関数として見ると $\mu(\{\omega \mid X(\omega) = a\})$ のことである。「刻一刻」の刻が ω に対応していて、(X が刻一刻と値を変えるということを数学的に表現するのはなんといいかとても難しいのだが、) コルモゴロフはそこらへんに目をつけて静的に定義することに成功したんじゃないのだろうかと思う。

動的概念を静的概念に変換するのは「数学というフォーマット」に乗せるための手続き

でしかなく、動的概念としての定義付けもそれはそれとして (数学というフォーマットの
外側で) 有効であり続けると個人的には思う。(集合論の公理系 ZFC は現代数学を「実装」
することができる (だろう) けど、別に ZFC 全然知らなくても全然数学できるし数学の論
文全然書けるしというのと類似の事柄だと思う。)

5.2 確率変数の向きを逆に

数学的に確率変数とは確率空間 $(\Omega, \mathfrak{M}, \mu)$ 上の実数値可測関数 $X: \Omega \rightarrow \mathbb{R}$ である。多
分、双対圏やらを考えて矢印の向きが $X: \Omega \leftarrow \mathbb{R}$ と考えた方がいい気がする。なぜなら、
確率変数を考える際に Ω は結構どうでもいいからである。これは通常の数学において、写
像の値域は結構どうでもいいことに相似している。こういうこと言うとそんなことないと言
われそうだけど、写像 $f: X \rightarrow Y$ に対して $X \rightarrow Y \times Z; x \rightarrow (f(x), o)$ みたいなのを
考えるときにめんどくさいしおんなじ記号で書きちゃうことはよくあるのでこういうこと
です。実際、確率変数の方を適当に無限直積を取って独立な確率変数を得る方法とかは、
確率変数の取り扱いにおいて定義域の方がどうでもいいということをよく利用していると思
う。この矢印逆にすると言うのは先に述べたプリミティブな確率変数の定義にも合致す
ると思う。なぜなら X を刻一刻と勝手に値が変わる変数として見ると、 $X = a$ となる瞬
間を観察してそこから $X = a$ となる確率の方が大事だったりするので $a \in \mathbb{R}$ に対して
“ $X = a$ となる事象” を対応させているので X は $\mathbb{R}$ から Ω の方に矢印が伸びてると考え
た方が便利なのである。このことは $X^{-1}: \mathfrak{B}(\mathbb{R}) \rightarrow \mathfrak{M}$ の方が大事ということを指し示し
ていると見ることもできる。おそらく確率変数を可測関数として見出した人は、“ $X = a$
という事象” の全体を抽象的に確率空間 Ω として用意してあげて、確率変数 X を改めて
可測関数 $X: \Omega \rightarrow \mathbb{R}$ というふうに見立てて、数学的に、静的に、確率変数を定義するこ
とに成功したのだと思う。確率変数の定義域は結構どうでもいいので、フワーっと Ω を
こさえてやっても適宜直積取ったりなんなりすればいいのでこの定義はうまく行ったと思
う (この定義にしたから定義域がどうでもよくなったのか、定義域がどうでも良いからこ
の定義になったのかはよくわからない)。ただし動的概念を扱う人と静的概念を扱う人との
間で齟齬が生まれるようになった、というのが現代のこの世界である。携帯電話の良い
ところはいつでもどこでも電話をかけられるところで、悪いところはいつでもどこでも
電話がかかってくるところである (過程を省くが結論だけ述べると仲良くしなさいという
意味)。

参考文献

- [1] はてなブログ電波通信の記事「直積測度の構成について：確率論その 0」,
<https://conscious4410.hatenablog.com/entry/2022/01/03/011245>
- [2] はてなブログ電波通信の記事「大数の法則について：確率論その 1」,
<https://conscious4410.hatenablog.com/entry/2022/01/16/233527>
- [3] はてなブログ電波通信の記事「距離空間上の測度：確率論その 2」,
<https://conscious4410.hatenablog.com/entry/2023/06/20/232112>.
- [4] はてなブログの記事「特殊な積分を求めること」,
<https://conscious4410.hatenablog.com/entry/2021/03/25/100050>
- [5] 小谷眞一, 「測度と確率」, 2005 岩波書店.
- [6] 船木直久, 「確率論」, 講座・数学の考え方 20, 2004, 朝倉書店.
- [7] C. W. Chin, *A short elementary proof of the central limit theorem by individual swapping*, arXiv:2106.00871
- [8] K. R. Parthasarathy, *Probability measures on metric spaces*. Academic Press.

「知識は公共財」