2024/9/1

中心極限定理

野津

1. はじめに

独立同分布に従うn個の確率変数の和は$n \to \infty$のとき正規分布に従う．このことを中心極限定理という．中心極限定理の証明は以下に述べるようにフーリエ変換の応用として最も興味深いものの一つである．

2. 特性関数

一般に確率密度関数$p(x)$に従う確率変数Xに対し，$p(x)$を次式によりフーリエ変換することで得られるλの関数$\phi(\lambda)$を特性関数と呼ぶ．

$$\phi(\lambda) = \int_{-\infty}^{\infty} p(x) e^{-i\lambda x} dx \tag{1}$$

フーリエ変換には$e^{-i\lambda x}$を用いる流儀と$e^{+i\lambda x}$を用いる流儀があり，特性関数の定義には$e^{+i\lambda x}$を用いる文献が多いようであるが，ここでは地震工学の慣例に倣い$e^{-i\lambda x}$を用いる．特性関数が与えられると，それをフーリエ逆変換することで確率密度関数は一意に定まる．

3. 互いに独立な確率変数の和の特性関数

確率密度関数$p_1(x)$と$p_2(x)$に従う互いに独立な確率変数X_1とX_2があるとき，$X_1 + X_2$の確率密度関数を$p_3(x)$とすれば

$$p_3(x) = p_1(x) * p_2(x) \tag{2}$$

である．ここに$*$は合積を表す．これは次のように示すことができる．$X_1 + X_2$がx周辺の幅Δxの区間に入る確率（確率密度関数の定義により$p_3(x)\Delta x$）は，「X_1がξ周辺の幅$\Delta\xi$の区間に入り，かつ，X_2が$x - \xi$周辺の幅Δxの区間に入る」確率をすべての$\Delta\xi$について加算した値となることから，

$$p_3(x)\Delta x = \int_{-\infty}^{\infty} p_1(\xi) p_2(x - \xi) \Delta x \mathrm{d}\xi \tag{3}$$

である．よって

$$p_3(x) = \int_{-\infty}^{\infty} p_1(\xi) p_2(x - \xi) \mathrm{d}\xi = p_1(x) * p_2(x) \tag{4}$$

である．さて，フーリエ変換の性質により，合積のフーリエ変換はもとの関数のフーリエ変換の積であるから，X_1とX_2の特性関数を$\phi_1(\lambda)$と$\phi_2(\lambda)$，$X_1 + X_2$の特性関数を$\phi_3(\lambda)$とすれば

$$\phi_3(\lambda) = \phi_1(\lambda)\phi_2(\lambda) \tag{5}$$

である．

4. 中心極限定理の証明

$\{X_1, \ldots, X_n\}$を平均μ，分散σ^2の独立同分布に従う確率変数とする．このとき

$$Y_i = (X_i - \mu)/\sigma \tag{6}$$

により平均 0，分散 1 の新しい確率変数Y_iを定義し，Y_iの和をS_n，それを$\sqrt{n}$で除したものをZ_nとする．

$$S_n = \sum_{i=1}^{n} Y_i \tag{7}$$

$$Z_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i \tag{8}$$

S_nの確率密度関数を$p_S(s)$，特性関数を$\phi_S(\lambda)$，Z_nの確率密度関数を$p_Z(z)$，特性関数を$\phi_Z(\lambda)$とし，$p_Z(z)dz = p_S(s)ds$であることを利用すると

$$\phi_Z(\lambda) = \int_{-\infty}^{\infty} p_Z(z) e^{-i\lambda z} dz = \int_{-\infty}^{\infty} p_S(s) e^{-i\lambda z} ds = \int_{-\infty}^{\infty} p_S(s) e^{-i(\lambda/\sqrt{n})s} ds = \phi_S(\lambda/\sqrt{n}) \tag{9}$$

であり，さらに式(4)より

$$\phi_Z(\lambda) = \left(\phi_Y(\lambda/\sqrt{n})\right)^n \tag{10}$$

である．ここに$\phi_Y(\lambda)$はY_iの特性関数である．以下，Y_iの確率密度関数を$p_Y(y)$とし，$\phi_Y(\lambda)$を$\lambda = 0$の周りでテイラー展開する．Y_iの平均が 0 であること，分散が 1 であることを考慮すると

$$\phi_Y(0) = \int_{-\infty}^{\infty} p_Y(y) dy = 1 \tag{11}$$

$$\phi'_Y(0) = -i \int_{-\infty}^{\infty} y p_Y(y) e^{-i\lambda y} dy \Big|_{\lambda=0} = -i \int_{-\infty}^{\infty} y p_Y(y) dy = 0 \tag{12}$$

$$\phi''_Y(0) = -\int_{-\infty}^{\infty} y^2 p_Y(y) e^{-i\lambda y} dy \Big|_{\lambda=0} = -\int_{-\infty}^{\infty} y^2 p_Y(y) dy = -1 \tag{13}$$

であるから

$$\phi_Y(\lambda) = 1 - \frac{\lambda^2}{2} + \cdots \tag{14}$$

すなわち

$$\phi_Y\left(\frac{\lambda}{\sqrt{n}}\right) = 1 - \frac{\lambda^2}{2n} + O\left(n^{-3/2}\right) \tag{15}$$

である．これと

$$e^x = \lim_{n \to \infty} \left(1 + \frac{x}{n}\right)^n \tag{16}$$

を用いると，$n \to \infty$のとき

$$\phi_Z(\lambda) = \left(\phi_Y\left(\lambda/\sqrt{n}\right)\right)^n = \left(1 - \frac{\lambda^2}{2n} + O\left(n^{-3/2}\right)\right)^n = \left(1 + \frac{-\lambda^2/2 + O\left(n^{-1/2}\right)}{n}\right)^n \to e^{-\lambda^2/2} \tag{17}$$

である．式(17)の右辺は平均 0，分散 1 である正規分布の確率密度関数のフーリエ変換である（フーリエ変換はなぜ元に戻るのか？の式(4)参照）．したがって，nが十分に大きいとき，Z_nは$N(0,1)$に従い，Y_iの和は$N(0,n)$に従い，X_iの和は$N(n\mu, n\sigma^2)$に従う．■