確率統計基礎講義 A

若木宏文

wakaki@hiroshima-u.ac.jp
http://home.hiroshima-u.ac.jp/ wakaki/lecture/index.shtml

2019.4.17

目 次

誤差分布

単回帰モデル

$$y_i = \alpha + \beta x_i + \varepsilon_i \ (i = 1, \ldots, N)$$

y_i : 目的変数
x_i : 説明変数
ε_i : 誤差

1. $\mathrm{E}[\varepsilon_i] = 0$
2. $\mathrm{Var}[\varepsilon_i] = \mathrm{E}[\varepsilon_i^2] = \sigma^2$
3. $i \neq j \ \Rightarrow \ \mathrm{Cov}[\varepsilon_i, \varepsilon_j] = \mathrm{E}[\varepsilon_i \varepsilon_j] = 0$

α, β, σ^2 に関する推測のためには, 誤差 ε_i の分布が必要

誤差の分布として, 裾が重い（軽い）ものを利用したい
$\Rightarrow$ 混合正規分布, t-分布
誤差の分布として, 非対称なものを利用したい
$\Rightarrow$ 歪正規分布

1.2. 混合正規分布

定義

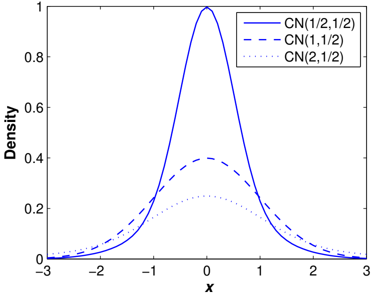
確率密度関数

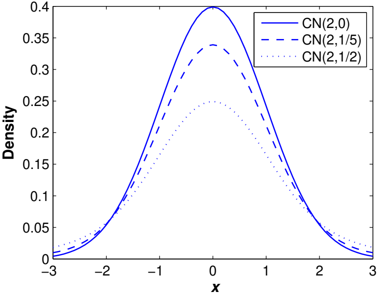
$$f(x;\lambda,\alpha) = (1-\alpha)\phi(x) + \frac{\alpha}{\lambda}\phi(x/\lambda)\ (\lambda > 0,\ \alpha \in [0,1])$$

によって定まる連続型分布を，1 次元混合正規分布といい，$\mathrm{CN}(\lambda,\alpha)$ と表す．
分散の異なる正規分布を混合比 α で混ぜた分布になっている．

分布の特性

1. 混合正規分布は 0 を中心に左右対称な分布であり, $\alpha = 0, 1$ か $\lambda = 1$ のときは正規分布となる.





分布の特性（続き）

2. $X \sim$ CN(λ, α) のとき,

$$\mathrm{E}[X] = 0,\ \mathrm{Var}[X] = 1 - \alpha + \alpha\lambda^2,$$
$$\kappa_3 = 0,\ \kappa_4 = \frac{3\alpha(1-\alpha)(\lambda^2-1)^2}{(1-\alpha+\alpha\lambda^2)^2}$$

平均周りのモーメント

$$\alpha_r = \begin{cases} 0 & (r \text{ が奇数}) \\ 1 \times 3 \times \cdots \times (r-1)(1-\alpha+\alpha\lambda^r) & (r \text{ が偶数}) \end{cases}$$

$\alpha = (1+\lambda^2)^{-1}$ とおくと,

$$f(x;\lambda) = \frac{\lambda^2}{1+\lambda^2}\phi(x) + \frac{1}{\lambda(1+\lambda^2)}\phi(x/\lambda)$$

となる. さらに, $\lambda = 1$ ならば標準正規分布である.
$\lambda \to \infty$ とすると, $\kappa_4 \to 0, X \xrightarrow{d} N(0,1)$ となる.

分布の特性（続き）

3. $X \sim \mathrm{CN}(\lambda, \alpha)$ の特性関数は

$$\mathcal{C}(t) = (1-\alpha)\exp\left(-\frac{1}{2}t^2\right) + \alpha\exp\left(-\frac{1}{2}t^2\lambda^2\right) \ (t \in \mathbb{R})$$

4. 乱数発生法：$Z_1, Z_2 \overset{i.i.d.}{\sim} N(0,1)$, $U \sim U(0,1)$ とし，Z_1, Z_2, U は独立とする．

$$X = \left\{1 - 1_{\{U \leq \alpha\}}\right\} Z_1 + 1_{\{U \leq \alpha\}} \lambda Z_2,$$
$$1_{\{U \leq \alpha\}} = \begin{cases} 1 & (U \leq \alpha) \\ 0 & (U > \alpha) \end{cases}$$

とすると，$X \sim \mathrm{CN}(\lambda, \alpha)$ となる．

レポート問題 1

混合正規分布に従う確率変数の, 平均, 分散, 歪度, 尖度, 平均周りの r 次モーメントを計算せよ. ただし, 正規分布の特性は既知であるとして計算して良い.

1.3. 歪正規分布

定義

確率密度関数

$$f(x;\lambda) = 2\phi(x)\Phi(\lambda x)\ (\lambda \in \mathbb{R})$$

によって定まる連続型分布を, 歪正規分布といい, SN(λ) と表す.
$f(x;\lambda)$ が確率密度関数になっていることは, 以下の定義により示される.
定理 **1.3.1** $f(x)$ を確率密度関数, $F(x)$ をその分布関数とする. f が偶関数であれば $2f(x)F(\lambda x)$ は確率密度関数となる.

畳み込み

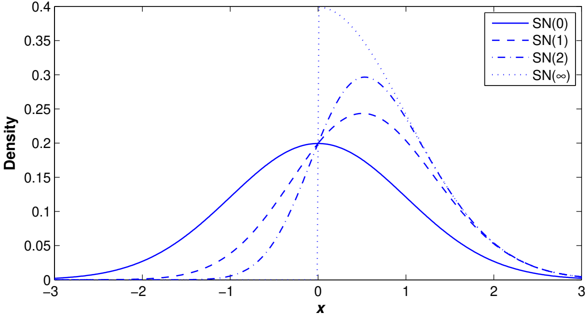
定理 1.3.2

$X \sim f(x)$（確率密度関数），$Y \sim g(y)$（確率密度関数），X, Y は独立とする．このとき，$Z = X + Y$ の確率密度関数 $h(z)$ は

$$h(z) = \int_{-\infty}^{\infty} f(z-y)g(y)dy$$

分布の特性

1. 歪正規分布は一般に非対称な分布であり, $\lambda = 0$ のときは, 正規分布となる. また, $\lambda = \to \pm\infty$ のときは半正規分布となる.



分布の特性（続き）

2. $X \sim$ SN(λ) のとき, $\delta = \dfrac{\lambda}{\sqrt{1+\lambda^2}}$ とおくと

$$\mathrm{E}[X] = \sqrt{\frac{2}{\pi}}\delta,\ \mathrm{Var}[X] = 1 - \frac{2}{\pi}\delta^2$$

$$\kappa_3 = \sqrt{2}(4-\pi)\mathrm{sign}(\lambda)\left\{\frac{\lambda^2}{\pi+(\pi-2)\lambda^2}\right\}^{3/2},$$

$$\kappa_4 = 8(\pi-3)\left\{\frac{\lambda^2}{\pi+(\pi-2)\lambda^2}\right\}^{2},$$

$$\lim_{\lambda\to\infty}\kappa_3 = \frac{\sqrt{2}(4-\pi)}{(\pi-2)^{3/2}},\ \lim_{\lambda\to\infty}\kappa_4 = \frac{8(\pi-3)}{(\pi-2)^2}$$

λ を大きくしても, 歪度や尖度はそれほど大きくならない.

分布の特性（続き）

3. $X \sim$ SN(λ) の特性関数は

$$\mathcal{C}(t) = \exp\left(-\frac{1}{2}t^2\right)\left\{1 + i\sqrt{\frac{2}{\pi}}\int_0^{t\lambda/\sqrt{1+\lambda^2}} e^{u^2/2}du\right\} \ (t \in \mathbb{R})$$

4. 乱数発生法：$Z_1, Z_2 \overset{i.i.d.}{\sim} N(0,1)$,

$$X = \frac{\mathrm{sign}(\lambda)}{\sqrt{1+\lambda^2}}(Z_1 + |\lambda Z_2|)$$

とすると, $X \sim$ SN(λ) となる.

レポート問題2

歪正規分布に従う確率変数の平均と分散を計算せよ.