

アクチュアリー「数学」演習

杉浦 誠

最終変更日: 2017年1月23日

目次

1	回帰分析	1
1.1	回帰直線 (単回帰)	1
1.2	重回帰	2
1.3	非線形回帰	3
1.4	確率分布の前提を置いた回帰モデルの分析	4
1.5	統計の復習 1 正規母集団と二項母集団	7
2	時系列解析	8
2.1	時系列に現れる確率過程と用語の定義	8
2.2	AR(p) (p 次の自己回帰モデル, Auto-regressive Model)	9
2.3	MA(q) (q 次の移動平均モデル, Moving-average Model)	11
2.4	ARMA(p, q) (Autoregressive Moving-average Model)	12
2.5	時系列モデルに基づく予測	13
2.6	統計の復習 2 順序統計量	14
3	確率過程	17
3.1	マルコフ連鎖とマルチンゲール	17
3.2	ポアソン過程	20
3.3	ブラウン運動	21
4	シミュレーション	22
4.1	確率変数を生成する技法	22
4.2	分散減少法	26
4.3	統計の復習 3 適合度、独立性の検定	29
5	損保数理に関する確率統計の話題から	32
5.1	最尤推定量の漸近挙動	32
5.2	極値問題	38
5.3	安定分布	45

これは 2014 年度後期に情報理論 II として行うアクチュアリー試験「数学」用の講義ノートです。教科書・参考書として以下を用いています。

- 日本アクチュアリー会編 モデリング 日本アクチュアリー会
- 藤田岳彦 著 確率・統計・モデリング問題集 日本アクチュアリー会
- 藤田岳彦 著 弱点克服大学生の確率・統計 東京図書, 2010
- 黒田耕嗣 著 生保年金数理 培風館, 2007
- 岩沢宏和 黒田耕嗣 著 損害保険数理 (アクチュアリー数学シリーズ 4), 日本評論社, 2015
- 国沢清典編 確率統計演習 2 統計 培風館, 1966
- 稲垣宣生 著 数理統計学 改訂版 裳華房, 2003
- 小寺平治 著 明解演習 数理統計 共立出版, 1986
- E.L. Lehmann, George Casella: Theory of Point Estimation, Second Edition, Springer, 1998
- S.I. Resnick: Extreme Values, Regular Variation and Point Processes, Springer, 1987
- 高橋 倫也, 志村 隆彰: 極値統計学 (ISM シリーズ:進化する統計数理), 近代科学社, 2016
- Breiman, L.: Probability, Addison-Wesley, 1968. (Classics in applied mathematics, 7, Society for Industrial and Applied Mathematics, 1992. Reprint 版)
- Durrett, R.: Probability Theory and Examples, 4th ed., Cambridge University Press, 2010.

教科書・参考書は今後増えていく予定です。

- 「藤田岳彦 著 確率・統計・モデリング問題集 日本アクチュアリー会」に従って述べていく。

1 回帰分析

1.1 回帰直線 (単回帰)

2 種類のデータの観測値 (x_i, y_i) , $(i = 1, 2, \dots, n)$ が与えられているとする。

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad (\text{データの平均})$$

$$s_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2 = \overline{x^2} - \bar{x}^2, \quad (\text{データの分散})$$

$$s_{xy} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \bar{y} = \overline{xy} - \bar{x} \bar{y}, \quad (\text{データの共分散})$$

$$r_{xy} = \frac{s_{xy}}{s_x s_y}, \quad (\text{データの相関係数}) \quad \text{ただし, } s_x = \sqrt{s_x^2}$$

などがデータの性質や関係を表す基本的な量である。以下の性質があった。

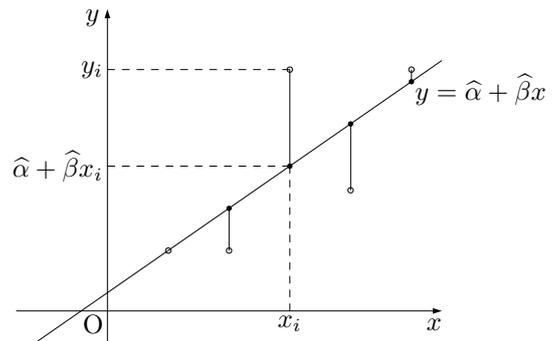
- $-1 \leq r_{xy} \leq 1$.
- $r_{xy} = 1$ (-1) \iff ある定数 $a > 0$ ($a < 0$) が存在し $\forall i$ に対して $y_i = ax_i + b$.
- $r_{xy} \equiv 1$ (-1) のとき、正の相関 (負の相関) が強いという。
- a, b, c, d を定数とし $ac > 0$ のとき、 $r_{ax+b, cy+d} = r_{xy}$. (相関係数は単位のとりかたによらない。)

• 最小二乗法

x_i から予測される値 $\alpha + \beta x_i$ と現実の値 y_i との差の二乗の和 $Q = \sum_{i=1}^n \{y_i - (\alpha + \beta x_i)\}^2$ が最小となるように係数 $\alpha = \hat{\alpha}, \beta = \hat{\beta}$ を定める:

$$0 = \frac{\partial Q}{\partial \alpha} = -2 \sum_{i=1}^n (y_i - (\hat{\alpha} + \hat{\beta} x_i)) = -2n(\bar{y} - \hat{\alpha} - \hat{\beta} \bar{x})$$

$$0 = \frac{\partial Q}{\partial \beta} = -2 \sum_{i=1}^n x_i (y_i - (\hat{\alpha} + \hat{\beta} x_i)) = -2n(\overline{xy} - \hat{\alpha} \bar{x} - \hat{\beta} \overline{x^2})$$



これより正規方程式

$$\begin{matrix} \hat{\alpha} + \hat{\beta} \bar{x} = \bar{y} \\ \hat{\alpha} \bar{x} + \hat{\beta} \overline{x^2} = \overline{xy} \end{matrix} \quad \text{行列表示で} \quad \begin{pmatrix} 1 & \bar{x} \\ \bar{x} & \overline{x^2} \end{pmatrix} \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \begin{pmatrix} \bar{y} \\ \overline{xy} \end{pmatrix} \quad (1.1)$$

これを解いて $\hat{\beta} = \frac{-\bar{x} \bar{y} + \overline{xy}}{\overline{x^2} - \bar{x}^2} = \frac{s_{xy}}{s_x^2} = r_{xy} \frac{s_y}{s_x}$, $\hat{\alpha} = \bar{y} - \hat{\beta} \bar{x} = \bar{y} - r_{xy} \frac{s_y}{s_x} \bar{x}$ を得る。

この $y = \hat{\alpha} + \hat{\beta} x$ を x を説明変数、 y を被説明変数とする回帰直線という。単に、 x から y への回帰直線ということもある。 $\hat{\alpha} = \bar{y} - \hat{\beta} \bar{x}$, $\hat{\beta} = r_{xy} \frac{s_y}{s_x}$ より、回帰直線は $y - \bar{y} = \hat{\beta}(x - \bar{x})$ あるいは

$$\frac{y - \bar{y}}{s_y} = r_{xy} \frac{x - \bar{x}}{s_x} \quad \text{つまり} \quad y \text{ の標準化} = \text{相関係数} \times x \text{ の標準化} \quad (1.2)$$

と表されることに注意する。また、回帰直線は (\bar{x}, \bar{y}) を通ることを注意する。

問題 1.1 x から y への回帰直線が $y = \alpha_1 + \beta_1 x$, y から x への回帰直線が $x = \alpha_2 + \beta_2 y$ であるとする。

- (1) $\beta_1 \beta_2 > 0$ のとき r_{xy} , s_y/s_x を β_1, β_2 を用いて表せ。
- (2) 更に、 $\beta_1 \beta_2 \neq 1$ と仮定する。 \bar{x}, \bar{y} を $\alpha_1, \alpha_2, \beta_1, \beta_2$ を用いて表せ。

• 決定係数

$\hat{y}_i = \hat{\alpha} + \hat{\beta}x_i$ を y_i の内挿値、 $e_i = y_i - \hat{y}_i$ を残差という。このとき、 $\sum_{i=1}^n e_i = 0$ 、 $\sum_{i=1}^n x_i e_i = 0$ が成立する。実際、 $\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$ と $\hat{\beta} = \frac{s_{xy}}{s_x^2}$ に注意すれば

$$\begin{aligned} \sum_{i=1}^n e_i &= \sum_{i=1}^n (y_i - \hat{y}_i) = n(\bar{y} - (\hat{\alpha} + \hat{\beta}\bar{x})) = 0, \\ \sum_{i=1}^n x_i e_i &= \sum_{i=1}^n x_i (y_i - \hat{y}_i) = n(\overline{xy} - (\hat{\alpha}\bar{x} + \hat{\beta}\bar{x}^2)) = n(\overline{xy} - (\bar{y} - \hat{\beta}\bar{x})\bar{x} - \hat{\beta}\bar{x}^2) = n(s_{xy} - \hat{\beta}s_x^2) = 0. \end{aligned}$$

全変動、回帰変動、残差変動について以下の関係式が成り立つ:

$$\begin{aligned} \text{全変動} &\equiv \sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (e_i + \hat{y}_i - \bar{y})^2 = \sum_{i=1}^n e_i^2 + 2 \sum_{i=1}^n e_i(\hat{y}_i - \bar{y}) + \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 \\ &= \sum_{i=1}^n e_i^2 + 2 \sum_{i=1}^n e_i(\hat{\alpha} + \hat{\beta}x_i - \bar{y}) + \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 \\ &= \sum_{i=1}^n e_i^2 + \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 \equiv \text{残差変動} + \text{回帰変動}. \end{aligned}$$

決定係数 $R^2 = 1 - \frac{\text{残差変動}}{\text{全変動}} = \frac{\text{回帰変動}}{\text{全変動}}$ と定める。 R^2 が 1 に近いほど回帰直線がデータによくあてはまっている。また、以下のように $R^2 = r_{xy}^2$ が示せる。

$$\begin{aligned} \text{全変動} &\equiv \sum_{i=1}^n (y_i - \bar{y})^2 = ns_y^2 \\ \text{回帰変動} &\equiv \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = \sum_{i=1}^n \left\{ \hat{\alpha} + \hat{\beta}x_i - (\hat{\alpha} + \hat{\beta}\bar{x}) \right\}^2 = \hat{\beta}^2 \sum_{i=1}^n (x_i - \bar{x})^2 = \left(r_{xy} \frac{s_y}{s_x} \right)^2 \cdot ns_x^2. \end{aligned}$$

問題 1.2 次に対し \bar{x} , \bar{y} , s_x^2 , s_y^2 , s_{xy} , $\hat{\beta}$, $\hat{\alpha}$ と、全変動、決定係数 R^2 , 回帰変動, 残差変動を求めよ。

(1) $(x_i, y_i) = (i, i^2)$ ($i = 1, 2, \dots, n$) ヒント: $\sum_{i=1}^n i^4 = \frac{n(n+1)(2n+1)(3n^2+3n-1)}{30}$ (導けるかな?)

i	1	2	3	4	5	6
x_i	1	2	2	4	5	5
y_i	5	14	11	21	18	26

i	1	2	3	4	5	6	7	8
x_i	3	6	8	9	6	7	3	5
y_i	4	7	8	9	4	5	5	6

(2) (3) は電卓を用い分数で表せ。)

1.2 重回帰

単回帰では説明変数が 1 つだったが、ここでは 2 個以上の場合を考える。簡単のため 2 個として説明する。

データの観測値 (x_{1i}, x_{2i}, y_i) , ($i = 1, 2, \dots, n$) が与えられているとし、 $Q = \sum_{i=1}^n \{y_i - (\alpha + \beta_1 x_{1i} + \beta_2 x_{2i})\}^2$ が最小となる $\hat{\alpha}, \hat{\beta}_1, \hat{\beta}_2$ は、 $\frac{\partial Q}{\partial \alpha} = \frac{\partial Q}{\partial \beta_1} = \frac{\partial Q}{\partial \beta_2} = 0$ より正規方程式

$$\begin{pmatrix} 1 & \bar{x}_1 & \bar{x}_2 \\ \bar{x}_1 & \bar{x}_1^2 & \bar{x}_1 \bar{x}_2 \\ \bar{x}_2 & \bar{x}_1 \bar{x}_2 & \bar{x}_2^2 \end{pmatrix} \begin{pmatrix} \hat{\alpha} \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} \bar{y} \\ \bar{x}_1 \bar{y} \\ \bar{x}_2 \bar{y} \end{pmatrix} \text{ を解いたもの}$$

となる。次のように書けることに注意する。 $(X^T$ は行列 X の転置を表す。)

$$X = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \end{pmatrix} \text{ とすると } XX^T = n \begin{pmatrix} 1 & \bar{x}_1 & \bar{x}_2 \\ \bar{x}_1 & \bar{x}_1^2 & \bar{x}_1 \bar{x}_2 \\ \bar{x}_2 & \bar{x}_1 \bar{x}_2 & \bar{x}_2^2 \end{pmatrix}, \quad X \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = n \begin{pmatrix} \bar{y} \\ \bar{x}_1 \bar{y} \\ \bar{x}_2 \bar{y} \end{pmatrix}.$$

また、以下が成り立つ。

回帰式 $y = \hat{\alpha} + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2$ は $(\bar{x}_1, \bar{x}_2, \bar{y})$ を通る。

残差 $e_i = y_i - \hat{y}_i = y_i - (\hat{\alpha} + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i})$ について $\sum_{i=1}^n e_i = 0, \sum_{i=1}^n x_{1i} e_i = 0, \sum_{i=1}^n x_{2i} e_i = 0$ が成立。

自由度修正決定係数 $\bar{R}^2 = 1 - \frac{\text{残差変動}/(n-k-1)}{\text{総変動}/(n-1)}$ 。

ここで、残差変動 $= \sum_{i=1}^n e_i^2$, 総変動 $= \sum_{i=1}^n (y_i - \bar{y})^2$, n は観測値の数、 k は説明変数の数である。

問題 1.3 五個のデータ $(x_{11}, y_{21}, y_1), \dots, (x_{15}, y_{25}, y_5)$ が与えられている。ここで、 $\sum x_{1i} = 3, \sum x_{2i} = 2, \sum y_i = 5, \sum x_{1i} x_{2i} = 4, \sum x_{1i} y_i = 12, \sum x_{2i} y_i = 8, \sum x_{1i}^2 = 10, \sum x_{2i}^2 = 12, \sum y_i^2 = 16$ であった。 y を x_1, x_2 で線形回帰するときの回帰式を求めよ。

● ダミー変数

データ (x_i, y_i) , $(i = 1, 2, \dots, n)$ から、奇数時点と偶数時点で定数項 α を変えた回帰式 $y = \alpha + \beta x$ を考える。

$$\text{ダミー変数 } d_i = \begin{cases} 1 & i = \text{奇数} \\ 0 & i = \text{偶数} \end{cases}$$

として、データ (x_i, d_i, y_i) から、回帰式 $y = \hat{\alpha} + \hat{\beta}_1 d + \hat{\beta}_2 x$ を考えると、

奇数時点では回帰式 $y = \hat{\alpha} + \hat{\beta}_1 + \hat{\beta}_2 x$, 偶数時点では回帰式 $y = \hat{\alpha} + \hat{\beta}_2 x$ と、定数項のみを変えた回帰式を求めることができる。

同様に、奇数時点と偶数時点で係数 β を変えた回帰式 $y = \alpha + \beta x$ は、上記のダミー変数を用い、

データ $(x_i, d_i x_i, y_i)$ から、回帰式 $y = \hat{\alpha} + \hat{\beta}_1 x + \hat{\beta}_2 (dx)$ を考えると、

奇数時点では回帰式 $y = \hat{\alpha} + (\hat{\beta}_1 + \hat{\beta}_2)x$, 偶数時点では回帰式 $y = \hat{\alpha} + \hat{\beta}_1 x$ と求めることができる。

問題 1.4 (1) 問題 1.2 (2) のデータについて、定数項ダミーを用いて奇数時点と偶数時点で定数項 α を変えた回帰式 $y = \alpha + \beta x$ を推定せよ。

(2) 問題 1.2 (2) のデータについて、係数ダミーを用いて奇数時点と偶数時点で係数 β を変えた回帰式 $y = \alpha + \beta x$ を推定せよ。

(3) 問題 1.2 (3) のデータの前半 4 つ上半期、後半 4 つは下半期についてであった。適当な定数項ダミー d を入れることにより、上半期と下半期で定数項 α を変えた回帰式 $y = \alpha + \beta x$ を推定せよ。

1.3 非線形回帰

あるタイプの非線形関数で当てはめるべきケースがある。ここでは応用上よく用いられるものを紹介する。

対数線形モデル $y = \alpha x^\beta$ の両辺の対数をとると、 $\log y = \log \alpha + \beta \log x$ 。

新しい変数として $y' = \log y, x' = \log x$ をとるとよい。

指数関数モデル $y = \alpha e^{\beta x}$ は、 $\log y = \log \alpha + \beta x$ と変形せよ。

変数 y のみを $y' = \log y$ に変える。

ロジスティック関数モデル $y = \frac{e^{\alpha + \beta x}}{1 + e^{\alpha + \beta x}}$ ($\beta > 0$) これは微分方程式 $\frac{dy}{dx} = \beta y(1 - y)$ ($0 < y < 1$) の解

変数 y のみを $y' = \log \frac{y}{1-y}$ に変える。

2 項回帰モデル 発生確率 y ($0 \leq y \leq 1$) が説明変数に依存して決まる回帰モデル。

これをある確率分布の分布関数 F を用いて $y = F(\alpha + \beta x)$ と表すと、 $y' = F^{-1}(y)$ とおくと $y' = \alpha + \beta x$ 。

$F(x) = \Phi(x)$ が $N(0, 1)$ の分布関数のとき、**プロビット・モデル**という。

$F(x) = \frac{e^x}{1 + e^x}$ (ロジスティック分布の分布関数) のとき、**ロジット・モデル**という。これはロジスティック関数モデルと同一のものである。

問題 1.5 (x, y) のデータが表のとおりに与えられている。このデータから、ロジット・モデル

$y = \frac{e^{\alpha+\beta x}}{1 + e^{\alpha+\beta x}}$ ($\beta > 0$) を用いた回帰式を求めるとき、 α, β の値を求めよ。また、プロビット・モデルの場合も求めよ。ただし小数点以下第 2 位まで求めよ。

x	1.2	1.4	2.7	3.5	4.8
y	10%	10%	50%	80%	90%

1.4 確率分布の前提を置いた回帰モデルの分析

1.4.1 推定量の分布

説明変数を x_i 、誤差項を確率変数 ε_i とし、被説明変数 Y_i を

$$Y_i = \alpha + \beta x_i + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (1.3)$$

とし、次を仮定する。

$$\varepsilon_1, \dots, \varepsilon_n \text{ は独立で } \varepsilon_i \sim N(0, \sigma^2). \quad (1.4)$$

このとき、最小二乗推定量 $\hat{\alpha}, \hat{\beta}$ および誤差項の分散 $\hat{\sigma}^2$ を考える:

$$\hat{\beta} = \frac{s_{xY}}{s_x^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})(Y_i - \bar{Y})}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad \hat{\alpha} = \bar{Y} - \hat{\beta}\bar{x}, \quad \hat{\sigma}^2 = \frac{1}{n-2} \sum_{i=1}^n (Y_i - (\hat{\alpha} + \hat{\beta}x_i))^2. \quad (1.5)$$

定理 1.1 仮定 (1.4) の下、 $\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2$ について以下が成立する。

$$(1) \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} \sim N\left(\begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \sigma^2 \begin{pmatrix} \frac{1}{n} + \frac{\bar{x}^2}{ns_x^2} & -\frac{\bar{x}}{ns_x^2} \\ -\frac{\bar{x}}{ns_x^2} & \frac{1}{ns_x^2} \end{pmatrix}\right) = N\left(\begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \frac{\sigma^2}{n} \begin{pmatrix} 1 & \bar{x} \\ \bar{x} & \bar{x}^2 \end{pmatrix}^{-1}\right).$$

$$\text{特に、} \hat{\alpha} \sim N\left(\alpha, \sigma^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{ns_x^2}\right)\right), \hat{\beta} \sim N\left(\beta, \frac{\sigma^2}{ns_x^2}\right).$$

$$(2) \frac{(n-2)\hat{\sigma}^2}{\sigma^2} \sim \chi_{n-2}^2. \text{ ただし、} \chi_{n-2}^2 \text{ は自由度 } n-2 \text{ のカイ二乗分布を表す。}$$

$$(3) \hat{\alpha} \text{ と } \hat{\sigma}^2 \text{ は独立。また、} \hat{\beta} \text{ と } \hat{\sigma}^2 \text{ は独立。}$$

証明: $\hat{\alpha}, \hat{\beta}$ を $\varepsilon_1, \dots, \varepsilon_n$ の線形結合で表す。 $c_i = \frac{x_i - \bar{x}}{\sum_{i=1}^n (x_i - \bar{x})^2}$ とし、 $\sum_{i=1}^n c_i = 0$ および

$$\sum_{i=1}^n c_i x_i = \frac{1}{ns_x^2} \left(\sum_{i=1}^n x_i^2 - \bar{x} \sum_{i=1}^n x_i \right) = \frac{\bar{x}^2 - \bar{x}^2}{s_x^2} = 1$$

に注意すると、

$$\hat{\beta} = \frac{\sum_{i=1}^n (x_i - \bar{x})(Y_i - \bar{Y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \sum_{i=1}^n c_i (\alpha + \beta x_i + \varepsilon_i) - \sum_{i=1}^n c_i \bar{Y} = \beta + \sum_{i=1}^n c_i \varepsilon_i,$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{x} = \frac{1}{n} \sum_{i=1}^n (\alpha + \beta x_i + \varepsilon_i) - \left(\beta + \sum_{i=1}^n c_i \varepsilon_i \right) \bar{x} = \alpha + \sum_{i=1}^n \left(\frac{1}{n} - c_i \bar{x} \right) \varepsilon_i.$$

従って、(1.4) より $(\hat{\alpha}, \hat{\beta})$ は二次元正規分布に従い、 $E[\hat{\alpha}] = \alpha, E[\hat{\beta}] = \beta$ 。また、 $\sum_{i=1}^n c_i^2 = \frac{ns_x^2}{(ns_x^2)^2} = \frac{1}{ns_x^2}$ に注意して、

$$V(\hat{\beta}) = \sum_{i=1}^n c_i^2 V(\varepsilon_i) = \frac{1}{ns_x^2} \sigma^2,$$

$$V(\hat{\alpha}) = \sum_{i=1}^n \left(\frac{1}{n} - c_i \bar{x} \right)^2 V(\varepsilon_i) = \sum_{i=1}^n \left(\frac{1}{n^2} - \frac{2c_i \bar{x}}{n} + c_i^2 \bar{x}^2 \right) \sigma^2 = \left(\frac{1}{n} + \frac{\bar{x}^2}{ns_x^2} \right) \sigma^2,$$

$$\text{Cov}(\hat{\alpha}, \hat{\beta}) = \sum_{i=1}^n \left(\frac{1}{n} - c_i \bar{x} \right) c_i V(\varepsilon_i) = -\frac{\bar{x}}{ns_x^2} \sigma^2$$

より (1) の \sim は示せる。最後の等号は $1 + \frac{\bar{x}^2}{s_x^2} = \frac{\bar{x}^2}{s_x^2}$ に注意して逆行列を計算せよ。

(2) $\hat{\sigma}^2$ は $e_i = Y_i - (\hat{\alpha} + \hat{\beta}x_i)$ とすると、 $\sum_{i=1}^n e_i = 0$, $\sum_{i=1}^n c_i e_i = 0$ の二つの制約条件があるため、自由度が 2 つ減って χ_{n-2}^2 に従うと説明される。(3) のため厳密な証明の概略を述べる: 1 行目が $(1/\sqrt{n}, \dots, 1/\sqrt{n})$,

2 行目が $\left(\frac{x_1 - \bar{x}}{\sqrt{ns_x^2}}, \dots, \frac{x_n - \bar{x}}{\sqrt{ns_x^2}} \right)$ で与えられる直行行列を A とする。このとき、 $\begin{pmatrix} \tilde{\varepsilon}_1 \\ \vdots \\ \tilde{\varepsilon}_n \end{pmatrix} = A \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}$ と定めると

$\tilde{\varepsilon}_1, \dots, \tilde{\varepsilon}_n$ は独立で $\tilde{\varepsilon}_i \sim N(0, \sigma^2)$ となる。また、 $\tilde{\varepsilon}_1 = \frac{1}{\sqrt{n}} \sum_{i=1}^n \varepsilon_i$, $\tilde{\varepsilon}_2 = \sum_{i=1}^n \frac{x_i - \bar{x}}{\sqrt{ns_x^2}} \varepsilon_i$, $\sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n \tilde{\varepsilon}_i^2$ より

$$\hat{\alpha} - \alpha = \frac{1}{\sqrt{n}} \tilde{\varepsilon}_1 - \frac{\bar{x}}{\sqrt{ns_x^2}} \tilde{\varepsilon}_2, \quad \hat{\beta} - \beta = \frac{1}{\sqrt{ns_x^2}} \tilde{\varepsilon}_2,$$

$$(n-2)\hat{\sigma}^2 = \sum_{i=1}^n \left\{ \varepsilon_i - (\hat{\alpha} - \alpha) - (\hat{\beta} - \beta)x_i \right\}^2 = \sum_{i=1}^n \left\{ \varepsilon_i - \frac{1}{\sqrt{n}} \tilde{\varepsilon}_1 - \frac{x_i - \bar{x}}{\sqrt{ns_x^2}} \tilde{\varepsilon}_2 \right\}^2 = \dots = \sum_{i=3}^n \tilde{\varepsilon}_i^2.$$

よって (2) は明らか。また、 $\hat{\alpha}, \hat{\beta}$ は $\tilde{\varepsilon}_1, \tilde{\varepsilon}_2$ の、 $\hat{\sigma}^2$ は $\tilde{\varepsilon}_3, \dots, \tilde{\varepsilon}_n$ の関数なので (3) も従う。 \square

1.4.2 α, β の区間推定と検定

自由度 n の t 分布は独立な $Z \sim N(0, 1)$ と $Y \sim \chi_n^2$ を用いて、 $T = \frac{Z}{\sqrt{Y/n}}$ の分布と定義されることに注意する。また、 $t_n(\alpha)$ で自由度 n の t 分布の上側 α 点: $T \sim t_n$ のとき $P(T \geq t_n(\alpha)) = \alpha$ とする。

以下、定理 1.1 を引用なしに頻繁に用いる。

- α の信頼区間: $Z = \frac{\hat{\alpha} - \alpha}{\sqrt{V(\hat{\alpha})}} \sim N(0, 1)$, $Y = \frac{(n-2)\hat{\sigma}^2}{\sigma^2} \sim \chi_{n-2}^2$ で Z と Y は独立なので、

$$T = \frac{Z}{\sqrt{Y/(n-2)}} = \frac{\hat{\alpha} - \alpha}{\sqrt{V(\hat{\alpha})\hat{\sigma}^2/\sigma^2}} = \frac{\hat{\alpha} - \alpha}{\sqrt{\hat{\sigma}^2\left(\frac{1}{n} + \frac{\bar{x}^2}{ns_x^2}\right)}} \sim t_{n-2}.$$

従って、信頼度 $1 - \varepsilon$ での α の信頼区間は

$$\hat{\alpha} - t_{n-2}(\varepsilon/2) \sqrt{\hat{\sigma}^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{ns_x^2} \right)} \leq \alpha \leq \hat{\alpha} + t_{n-2}(\varepsilon/2) \sqrt{\hat{\sigma}^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{ns_x^2} \right)}.$$

- β の信頼区間: $Z = \frac{\hat{\beta} - \beta}{\sqrt{V(\hat{\beta})}} \sim N(0, 1)$, $Y = \frac{(n-2)\hat{\sigma}^2}{\sigma^2} \sim \chi_{n-2}^2$ で Z と Y は独立なので、

$$T = \frac{Z}{\sqrt{Y/(n-2)}} = \frac{\hat{\beta} - \beta}{\sqrt{V(\hat{\beta})\hat{\sigma}^2/\sigma^2}} = \frac{\hat{\beta} - \beta}{\sqrt{\hat{\sigma}^2 \frac{1}{ns_x^2}}} \sim t_{n-2}.$$

従って、信頼度 $1 - \varepsilon$ での β の信頼区間は

$$\hat{\beta} - t_{n-2}(\varepsilon/2) \sqrt{\frac{\hat{\sigma}^2}{ns_x^2}} \leq \beta \leq \hat{\beta} + t_{n-2}(\varepsilon/2) \sqrt{\frac{\hat{\sigma}^2}{ns_x^2}}.$$

注意 1.1 $\hat{\sigma}^2$ を計算する際は $\hat{\sigma}^2 = \frac{1}{n-2}(1 - r_{xy}^2)ns_y^2$ と計算するのがよい。これは次のように導かれる。

$$\hat{\sigma}^2 = \frac{1}{n-2} \sum_{i=1}^n e_i^2 = \frac{1}{n-2} (\text{残差変動}) = \frac{1}{n-2} (\text{全変動} - \text{回帰変動}) = \frac{1}{n-2} (1 - \text{決定係数 } R^2) (\text{全変動})$$

$$= \frac{1}{n-2}(1-r_{xy}^2)ns_y^2.$$

- **検定:** 次の手順で有意水準 ε の両側検定を行うことができる。

帰無仮説 $H_0: \beta = \beta_0$, 対立仮説 $H_1: \beta \neq \beta_0$

H_0 のもとで、 $T = \frac{\hat{\beta} - \beta_0}{\sqrt{\frac{\hat{\sigma}^2}{ns_x^2}}} \sim t_{n-2}$. よって、 t 分布表から $t_{n-2}(\varepsilon/2)$ を求め、標本からの実現値 t に対して、

$|t| > t_{n-2}(\varepsilon/2)$ なら H_0 を棄却、 $|t| \leq t_{n-2}(\varepsilon/2)$ なら H_0 を採択すればよい。

同様に、片側検定の場合、

帰無仮説 $H_0: \beta = \beta_0$, 対立仮説 $H_1: \beta > \beta_0$ のときは、

$t > t_{n-2}(\varepsilon)$ なら H_0 を棄却、 $t \leq t_{n-2}(\varepsilon)$ なら H_0 を採択すればよい。

帰無仮説 $H_0: \beta = \beta_0$, 対立仮説 $H_1: \beta < \beta_0$ のときは、

$t < t_{n-2}(\varepsilon)$ なら H_0 を棄却、 $t \geq t_{n-2}(\varepsilon)$ なら H_0 を採択すればよい。

問題 1.6 問題 1.2 (2) のデータについて、 $\hat{\sigma}^2$ の実現値を求め、 α, β の 95% 信頼区間を求めよ。ただし小数点以下第 3 位まで求めよ。

問題 1.7 問題 1.2 (3) のデータについて、帰無仮説 $H_0: \beta = 0$, 対立仮説 $H_1: \beta > 0$ を、有意水準 5% で検定せよ。

1.4.3 点予測、区間予測

説明変数 x_{n+1} が与えられたときの Y_{n+1} の予測量 \hat{Y}_{n+1} は、 $\hat{\alpha}, \hat{\beta}$ を用いて、 $\hat{Y}_{n+1} = \hat{\alpha} + \hat{\beta}x_{n+1}$ となり。これは正規分布に従う。(これは $\hat{\alpha}, \hat{\beta}$ が $\varepsilon_1, \dots, \varepsilon_n$ の線形結合であることによる。)

予測誤差 $Y_{n+1} - \hat{Y}_{n+1} = -(\hat{\alpha} - \alpha) - (\hat{\beta} - \beta)x_{n+1} + \varepsilon_{n+1}$ について定理 1.1(1) より、

$$\begin{aligned} E[Y_{n+1} - \hat{Y}_{n+1}] &= -E[\hat{\alpha} - \alpha] - x_{n+1}E[\hat{\beta} - \beta] + E[\varepsilon_{n+1}] = 0, \\ V(Y_{n+1} - \hat{Y}_{n+1}) &= V(\varepsilon_{n+1} - (\hat{\alpha} - \alpha) - (\hat{\beta} - \beta)x_{n+1}) = V(\varepsilon_{n+1} - \hat{\alpha} - \hat{\beta}x_{n+1}) \\ &= V(\varepsilon_{n+1}) + V(-\hat{\alpha} - \hat{\beta}x_{n+1}), \quad (\because \varepsilon_{n+1} \text{ は } \hat{\alpha}, \hat{\beta} \text{ と独立}) \\ &= \sigma^2 + V(\hat{\alpha}) + 2x_{n+1} \text{Cov}(\hat{\alpha}, \hat{\beta}) + x_{n+1}^2 V(\hat{\beta}) \\ &= \sigma^2 + \sigma^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{ns_x^2} \right) + 2x_{n+1} \frac{-\bar{x}\sigma^2}{ns_x^2} + x_{n+1}^2 \frac{\sigma^2}{ns_x^2} \\ &= \sigma^2 \left\{ 1 + \frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{ns_x^2} \right\} \end{aligned}$$

従って、予測誤差 $Y_{n+1} - \hat{Y}_{n+1} \sim N\left(0, \sigma^2 \left\{ 1 + \frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{ns_x^2} \right\}\right)$ となる。

これより、 σ^2 が既知であればこれより区間推定できる。

σ^2 が未知の場合、 $Z = \frac{Y_{n+1} - \hat{Y}_{n+1}}{\sqrt{V(Y_{n+1} - \hat{Y}_{n+1})}} \sim N(0, 1)$, $W = \frac{(n-2)\hat{\sigma}^2}{\sigma^2} \sim \chi_{n-2}^2$ で Z と W は独立なので、

$$T = \frac{Z}{\sqrt{W/(n-2)}} = \frac{Y_{n+1} - \hat{Y}_{n+1}}{\sqrt{V(Y_{n+1} - \hat{Y}_{n+1})\hat{\sigma}^2/\sigma^2}} = \frac{Y_{n+1} - \hat{Y}_{n+1}}{\sqrt{\hat{\sigma}^2 \left\{ 1 + \frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{ns_x^2} \right\}}} \sim t_{n-2}.$$

従って、信頼度 $1 - \varepsilon$ での Y_{n+1} の信頼区間は

$$\hat{Y}_{n+1} - t_{n-2}\left(\frac{\varepsilon}{2}\right) \sqrt{\hat{\sigma}^2 \left\{ 1 + \frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{ns_x^2} \right\}} \leq Y_{n+1} \leq \hat{Y}_{n+1} + t_{n-2}\left(\frac{\varepsilon}{2}\right) \sqrt{\hat{\sigma}^2 \left\{ 1 + \frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{ns_x^2} \right\}}$$

となる。ただし、上式で \hat{Y}_{n+1} と $\hat{\sigma}^2$ は実現値を表す。

問題 1.8 問題 1.2 (3) のデータに対して推定された回帰式を用いて (問題 1.7 も参照のこと)、 $x_{8+1} = 4$ に対する点予測および信頼係数 95% 信頼区間を求めよ。ただし小数点以下第 3 位まで求めよ。

1.5 統計の復習 1 正規母集団と二項母集団

定義 1.1 正規母集団の統計において次の分布は特に重要である。

χ^2 分布: X_1, \dots, X_n が i.i.d. で $N(0, 1)$ に従うとき、 $X_1^2 + \dots + X_n^2 \sim \chi_n^2$ (自由度 n の χ^2 分布)。

t 分布: Y, Z は独立で $Y \sim \chi_n^2, Z \sim N(0, 1)$ のとき、 $T = \frac{Z}{\sqrt{Y/n}} \sim t_n$ (自由度 n の t 分布)。

F 分布: X, Y は独立で $X \sim \chi_m^2, Y \sim \chi_n^2$ のとき、 $W = \frac{X/m}{Y/n} \sim F_n^m$ (自由度 (m, n) の F 分布)。

次の定理は確率統計学 I で定理 3.7 で示した。

定理 1.2 X_1, \dots, X_n が独立で、それぞれ同一の正規分布 $N(\mu, \sigma^2)$ に従うとすると、次が成立する。

(1) 標本平均 $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ は $N\left(\mu, \frac{\sigma^2}{n}\right)$ に従う。

(2) 不偏分散 $U^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ について、 $\frac{n-1}{\sigma^2} U^2 = \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2 \sim \chi_{n-1}^2$ 。

(3) \bar{X} と U^2 は独立。

応用例

• 正規母集団の母平均の区間推定、検定において、母分散 σ^2 が既知の場合、定理 1.2 (1) を用いて行うことができた。例えば、標本平均の実現値が \bar{x} のとき、信頼度 $1 - \varepsilon$ での母平均 μ の信頼区間は

$$\bar{x} - u(\varepsilon/2) \sqrt{\frac{\sigma^2}{n}} \leq \mu \leq \bar{x} + u(\varepsilon/2) \sqrt{\frac{\sigma^2}{n}},$$

ここで、 $u(\alpha)$ は $N(0, 1)$ の上側 α 点を表す。母分散が未知であっても、標本数が大きい場合は母分散を不偏分散の実現値としてこれを用いた。

• 正規母集団の母分散の区間推定、検定において、定理 1.2 (2) を用いて行うことができた。例えば、不偏分散の実現値が u^2 のとき、信頼度 $1 - \varepsilon$ での母分散 σ^2 の信頼区間は

$$\frac{(n-1)u^2}{\chi_{n-1}^2(\varepsilon/2)} \leq \sigma^2 \leq \frac{(n-1)u^2}{\chi_{n-1}^2(1-\varepsilon/2)},$$

ここで、 $\chi_{n-1}^2(\alpha)$ は χ_{n-1}^2 の上側 α 点を表す。

• 正規母集団の母平均の区間推定、検定において、母分散 σ^2 が未知の場合、定理 1.2 より $T = \frac{\bar{X} - \mu}{\sqrt{U^2/n}} \sim t_{n-1}$ となること (各自証明を試みよう) を用いて行うことができた。例えば、標本平均、不偏分散の実現値が \bar{x}, u^2 のとき、信頼度 $1 - \varepsilon$ での母平均 μ の信頼区間は

$$\bar{x} - t_{n-1}(\varepsilon/2) \sqrt{\frac{u^2}{n}} \leq \mu \leq \bar{x} + t_{n-1}(\varepsilon/2) \sqrt{\frac{u^2}{n}},$$

ここで、 $t_{n-1}(\alpha)$ は t_{n-1} の上側 α 点を表す。

• 2 つの正規母集団の母数の比較に関する検定を前期の数理統計学 I の最後の節で取り上げた。これは区間推定にも用いられる。例えば、

$X_1, \dots, X_m, Y_1, \dots, Y_n$ をそれぞれ正規母集団 $N(\mu_1, \sigma_1^2), N(\mu_2, \sigma_2^2)$ からの無作為標本とする。このとき、標本平均を \bar{X}, \bar{Y} 、標本分散を S_X^2, S_Y^2 とすると (母分散 σ_1^2 と σ_2^2 は既知であれば)、 $Z = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/m + \sigma_2^2/n}}$ は $N(0, 1)$ に従う。従って、標本平均の実現値を \bar{x}, \bar{y} とすると、平均の差 $\mu_1 - \mu_2$ の信頼度 $1 - \varepsilon$ の信頼区間は

$$\bar{x} - \bar{y} - u(\varepsilon/2) \sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}} \leq \mu_1 - \mu_2 \leq \bar{x} - \bar{y} + u(\varepsilon/2) \sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}},$$

ここで、母分散が未知であっても、標本数が大きい場合は母分散 σ_1^2, σ_2^2 をそれぞれの不偏分散の実現値 u_1^2, u_2^2 で置き換えて成立する。

• 上記は大標本での二項母集団の区間推定や検定にも用いることができる。

例えば、母比率 p_1 の二項母集団からの大きさ m の標本比率を \hat{P}_1 、母比率 p_2 の二項母集団からの大きさ n の標本比率を \hat{P}_2 とすると、二項分布の正規分布近似を考え、

$$\hat{P}_1 \sim N\left(p_1, \frac{p_1(1-p_1)}{m}\right), \hat{P}_2 \sim N\left(p_2, \frac{p_2(1-p_2)}{n}\right) \text{ より } \hat{P}_1 - \hat{P}_2 \sim N\left(p_1 - p_2, \frac{p_1(1-p_1)}{m} + \frac{p_2(1-p_2)}{n}\right).$$

これより、標準化し、標本比率と根号内の母比率をその実現値 \hat{p}_1, \hat{p}_2 に置き換えることで、母比率の差 $p_1 - p_2$ の信頼度 $1 - \varepsilon$ の信頼区間

$$\hat{p}_1 - \hat{p}_2 - u(\varepsilon/2) \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{m} + \frac{\hat{p}_2(1-\hat{p}_2)}{n}} \leq p_1 - p_2 \leq \hat{p}_1 - \hat{p}_2 + u(\varepsilon/2) \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{m} + \frac{\hat{p}_2(1-\hat{p}_2)}{n}}$$

を得る。

例題 1.1 ある政策の支持率を予想するために、母集団から男性 900 人、女性 800 人をそれぞれ無作為に抽出して調査を行ったところ、男性は 300 人、女性は 320 人が支持すると回答した。母集団全体の男女比は 5 : 4 であるとして、母集団全体での支持率を近似法を用いて、信頼度 95% で区間推定せよ。

解: 男女の支持率を p_1, p_2 とし、標本比率を \hat{P}_1, \hat{P}_2 とする。このとき、 $\hat{P}_1 \sim N\left(p_1, \frac{p_1(1-p_1)}{900}\right)$, $\hat{P}_2 \sim N\left(p_2, \frac{p_2(1-p_2)}{800}\right)$ と近似される。男女比を考慮すると全体の支持率は $\hat{P} = \frac{5}{9}\hat{P}_1 + \frac{4}{9}\hat{P}_2$ となるから、 \hat{P}_1 と \hat{P}_2 は独立なので、 $\hat{P} \sim N\left(\frac{5}{9}p_1 + \frac{4}{9}p_2, \left(\frac{5}{9}\right)^2 \frac{p_1(1-p_1)}{900} + \left(\frac{4}{9}\right)^2 \frac{p_2(1-p_2)}{800}\right)$ 。これより、標準化し、標本比率と根号内の母比率をその実現値 $\hat{p}_1 = \frac{300}{900}$, $\hat{p}_2 = \frac{320}{800}$ に置き換えることで、

$$\frac{5}{9}\hat{p}_1 + \frac{4}{9}\hat{p}_2 \pm u(0.025) \sqrt{\left(\frac{5}{9}\right)^2 \frac{\hat{p}_1(1-\hat{p}_1)}{900} + \left(\frac{4}{9}\right)^2 \frac{\hat{p}_2(1-\hat{p}_2)}{800}} = 0.36296 \dots \pm 0.02281 \dots = \begin{cases} 0.3857 \dots \\ 0.3401 \dots \end{cases}$$

従って、 $0.340 \leq p \leq 0.386$ 。 □

問題 1.9 ある都市の市長選挙の結果を予想するために、60 才未満の者 120 人に意見を求めたところ 48 人が保守系を支持すると言った。一方、60 才以上の者 80 人について調べたところ、56 人が保守系を支持した。

(1) 近似法を用いて、60 才以上の人の支持率と 60 才未満の人の支持率の差の信頼係数 95% 信頼区間を求めよ。ただし小数点以下第 3 位まで求めよ。

(2) 投票率を考慮すると、この都市の 60 才未満の人と 60 才以上の人の比は 4 : 5 である。近似法を用いて、この選挙での保守系の得票率を信頼度 95% で区間推定せよ。ただし小数点以下第 3 位まで求めよ。

2 時系列解析

2.1 時系列に現れる確率過程と用語の定義

確率過程 $Y_t, t = 0, \pm 1, \pm 2, \dots$ を考える。ここでは、時間パラメータ t は負の値もとることに注意する。

定義 2.1 確率過程 $Y_t, t = 0, \pm 1, \pm 2, \dots$ が定常 (stationary) であるとは、

条件 1 $E[Y_t] = \mu$ (定数)

条件 2 $\text{Cov}(Y_t, Y_{t-h}) = \gamma_h$, 特に、 $\text{Cov}(Y_t, Y_t) = V(Y_t) = \gamma_0$ (定数)

となるときにいう。

定義 2.2 $\gamma_h = \text{Cov}(Y_t, Y_{t-h})$ を時差 h の自己共分散という。

$$\rho_h = \frac{\gamma_h}{\gamma_0} = \frac{\text{Cov}(Y_t, Y_{t-h})}{V(Y_t)}$$
 を時差 h の自己相関 (コレログラム) という。

定常過程の典型例: $X_t, t = 0, \pm 1, \pm 2, \dots$ を i.i.d. とし、数列 $\{a_n\}$ が $\sum_{n=0}^{\infty} |a_n| < \infty$ を満たすとす。

$$Y_t = \sum_{i=0}^{\infty} a_i X_{t-i}$$
 とおくと、これは定常過程。実際、 $E[X_t] = \mu, V(X_t) = \sigma^2$ とすると、

$$\begin{aligned} E[Y_t] &= \sum_{i=0}^{\infty} a_i E[X_{t-i}] = \mu \sum_{i=0}^{\infty} a_i, \\ \text{Cov}(Y_t, Y_{t-h}) &= \text{Cov}\left(\sum_{i=0}^{\infty} a_i X_{t-i}, \sum_{j=0}^{\infty} a_j X_{t-h-j}\right) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} a_i a_j \text{Cov}(X_{t-i}, X_{t-h-j}) \\ &= \sigma^2 \sum_{i=0}^{\infty} a_{h+j} a_j, \quad (\because \text{Cov}(X_{t-i}, X_{t-h-j}) \neq 0 \iff i = h+j) \end{aligned}$$

定義 2.3 $\{y_t\}_{t=1}^n$ を $\{Y_t\}_{t=1}^n$ の実現値とする。このとき $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_t$ とする。

$$\hat{\gamma}_h = \frac{1}{n} \sum_{t=h+1}^n (y_t - \bar{y})(y_{t-h} - \bar{y})$$
 を時差 h の標本自己共分散という。

$$\hat{\rho}_h = \frac{\hat{\gamma}_h}{\hat{\gamma}_0} = \frac{\sum_{t=h+1}^n (y_t - \bar{y})(y_{t-h} - \bar{y})}{\sum_{t=1}^n (y_t - \bar{y})^2}$$
 を時差 h の標本自己相関という。

2.2 AR(p) (p 次の自己回帰モデル, Auto-regressive Model)

$\{Y_t\}$ が p 次の自己回帰モデル AR(p) に従うとは、

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t \quad (2.1)$$

となることをいう。ここで、 $\phi_0, \phi_1, \dots, \phi_p$ は定数であり、誤差項 ε_t は $Y_{t-1}, Y_{t-2}, \dots, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots$ と独立で $E[\varepsilon_t] = 0, V(\varepsilon_t) = \sigma^2$ とし、同一の分布に従う。 $\{Y_t\} \sim \text{AR}(p)$ と表す。

命題 2.1 (定常性) 特性方程式 $\phi(x) = 1 - (\phi_1 x + \phi_2 x^2 + \dots + \phi_p x^p) = 0$ の解^{*1}の絶対値がすべて 1 より大きいとき、AR(p) は定常性をもつ。

$p = 1$, つまり、 $Y_t = \phi_0 + \phi_1 Y_{t-1} + \varepsilon_t$ のとき、 $\phi(x) = 1 - \phi_1 x$ より、定常 $\iff |1/\phi_1| > 1 \iff |\phi_1| < 1$ 。このとき、(2.1) の t を $t-1, t-2, \dots$ として代入していくと、

$$\begin{aligned} Y_t &= \phi_0 + \phi_1 Y_{t-1} + \varepsilon_t = \phi_0 + \phi_1(\phi_0 + \phi_1 Y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\ &= \phi_0(1 + \phi_1) + \phi_1^2 Y_{t-2} + (\varepsilon_t + \phi_1 \varepsilon_{t-1}) = \phi_0(1 + \phi_1) + \phi_1^2(\phi_0 + \phi_1 Y_{t-3} + \varepsilon_{t-2}) + (\varepsilon_t + \phi_1 \varepsilon_{t-1}) \\ &= \phi_0(1 + \phi_1 + \phi_1^2) + \phi_1^3 Y_{t-3} + (\varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2}) \\ &= \dots \\ &= \phi_0(1 + \phi_1 + \dots + \phi_1^{n-1}) + \phi_1^n Y_{t-n} + (\varepsilon_t + \phi_1 \varepsilon_{t-1} + \dots + \phi_1^{n-1} \varepsilon_{t-n}). \end{aligned}$$

よって、 $|\phi_1| < 1$ に注意して $n \rightarrow \infty$ とすると次を得る。

$$Y_t = \frac{\phi_0}{1 - \phi_1} + \sum_{i=0}^{\infty} \phi_1^i \varepsilon_{t-i}. \quad \text{これは AR(1) の MA}(\infty) \text{ 表現である。} \quad (2.2)$$

$\{Y_t\} \sim \text{AR}(p)$ とし、(2.1) で期待値をとると、 $\mu = E[Y_t] = \phi_0 + \phi_1 \mu + \phi_2 \mu + \dots + \phi_p \mu$ となるので

$$Y_t - \mu = \phi_1(Y_{t-1} - \mu) + \dots + \phi_p(Y_{t-p} - \mu) + \varepsilon_t.$$

*1 ラグ作用素 $L: LY_t = Y_{t-1}$ を用いると、(2.1) は $Y_t = (\phi_1 L + \phi_2 L^2 + \dots + \phi_p L^p)Y_t + \phi_0 + \varepsilon_t$, 即ち $\phi(L)Y_t = \phi_0 + \varepsilon_t$ と表せる。

これに、 $Y_t - \mu, Y_{t-1} - \mu, \dots$ をかけて期待値をとれば、

$$\begin{aligned}\gamma_0 &= \phi_1 \gamma_1 + \phi_2 \gamma_2 + \dots + \phi_p \gamma_p + \sigma^2 \\ \gamma_1 &= \phi_1 \gamma_0 + \phi_2 \gamma_1 + \dots + \phi_p \gamma_{p-1} \\ \gamma_2 &= \phi_1 \gamma_1 + \phi_2 \gamma_0 + \dots + \phi_p \gamma_{p-2} \\ &\vdots \\ \gamma_p &= \phi_1 \gamma_{p-1} + \phi_2 \gamma_{p-2} + \dots + \phi_p \gamma_0 \\ \gamma_{p+1} &= \phi_1 \gamma_p + \phi_2 \gamma_{p-1} + \dots + \phi_p \gamma_1 \\ \gamma_{p+t} &= \phi_1 \gamma_{p+t-1} + \phi_2 \gamma_{p+t-2} + \dots + \phi_{p-1} \gamma_{t+1} + \phi_p \gamma_t \quad (t \geq 0)\end{aligned}$$

が成立する。逆に上式の両辺を γ_0 で割ることで、 $\mu, \gamma_0, \gamma_1, \dots, \gamma_p$ は、 $\phi_0 = (1 - (\phi_1 + \dots + \phi_p))\mu$ および

$$\begin{cases} \rho_1 = \phi_1 + \phi_2 \rho_1 + \dots + \phi_p \rho_{p-1} \\ \rho_2 = \phi_1 \rho_1 + \phi_2 + \dots + \phi_p \rho_{p-2} \\ \vdots \\ \rho_p = \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \dots + \phi_p \end{cases} \quad \text{行列で} \quad \begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{pmatrix} = \begin{pmatrix} 1 & \rho_1 & \dots & \rho_{p-1} \\ \rho_1 & 1 & \dots & \rho_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p-1} & \rho_{p-2} & \dots & 1 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{pmatrix} \quad (2.3)$$

を満たす。これをユール=ウォーカーの方程式という。 $\rho_h = \gamma_h / \gamma_0$ に注意。

- AR(1): $Y_t = \phi_0 + \phi_1 Y_{t-1} + \varepsilon_t$ のとき、 $\mu = \phi_0 + \phi_1 \mu$ より $\mu = \frac{\phi_0}{1 - \phi_1}$.

$$\begin{aligned}\gamma_0 &= \phi_1 \gamma_1 + \sigma^2, \gamma_1 = \phi_1 \gamma_0, \gamma_2 = \phi_1 \gamma_1, \dots \text{より} \\ \gamma_0 &= \frac{\sigma^2}{1 - \phi_1^2}, \gamma_1 = \frac{\phi_1 \sigma^2}{1 - \phi_1^2}. \text{ また、} \gamma_h = \frac{\phi_1^h \sigma^2}{1 - \phi_1^2}, \text{ 自己相関 } \rho_h = \phi_1^h.\end{aligned}$$

- AR(2): $Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t$ のとき、

特性方程式 $\phi(x) = 1 - \phi_1 x - \phi_2 x^2 = 0$ のすべての解の絶対値が 1 より大きい条件を求めると、

(i) 実数解をもつ場合 $D = \phi_1^2 + 4\phi_2 \geq 0, \phi(-1) = 1 + \phi_1 - \phi_2 > 0, \phi(1) = 1 - \phi_1 - \phi_2 > 0$.

(ii) 虚数解をもつ場合 $D = \phi_1^2 + 4\phi_2 < 0$,

$$\left| \frac{-\phi_1 \pm \sqrt{-(\phi_1^2 + 4\phi_2)}i}{2\phi_2} \right|^2 = \left(\frac{-\phi_1}{2\phi_2} \right)^2 + \left(\frac{\sqrt{-(\phi_1^2 + 4\phi_2)}}{2\phi_2} \right)^2 = \frac{\phi_1^2 + \{-(\phi_1^2 + 4\phi_2)\}}{4\phi_2^2} = \frac{-1}{\phi_2} > 1$$

$\phi_2 < 0$ より、 $\phi_2 > -1$ となる。

以上をまとめると、

$$\begin{aligned}\text{AR(2) が定常性をもつ} &\iff \phi_2 = 1 - |\phi_1| \text{ かつ } \phi_2 > -1 \\ &\iff (\phi_1, \phi_2) \text{ が } (0, 1), (-2, -1), (2, -1) \text{ を頂点とする三角形の内部}\end{aligned}$$

特に、放物線 $\phi_2 = -\frac{1}{4}\phi_1^2$ の上側が実数解、下側が虚数解に対応する。 (ϕ_1, ϕ_2) -平面に図示せよ。)

$$\text{平均: 期待値をとって、} \mu = \phi_0 + \phi_1 \mu + \phi_2 \mu + 0 \text{ より } \mu = \frac{\phi_0}{1 - (\phi_1 + \phi_2)}.$$

分散、自己共分散: $Y_t - \mu = \phi_1(Y_{t-1} - \mu) + \dots + \phi_2(Y_{t-2} - \mu) + \varepsilon_t$ より、

$$\begin{aligned}\gamma_0 &= \phi_1 \gamma_1 + \phi_2 \gamma_2 + \sigma^2, \\ \gamma_1 &= \phi_1 \gamma_0 + \phi_2 \gamma_1, \quad \gamma_1 = \frac{\phi_1}{1 - \phi_2} \gamma_0 \\ \gamma_2 &= \phi_1 \gamma_1 + \phi_2 \gamma_0, \quad \gamma_2 = \left(\frac{\phi_1^2}{1 - \phi_2} + \phi_2 \right) \gamma_0 \\ \gamma_n &= \phi_1 \gamma_{n-1} + \phi_2 \gamma_{n-2} \quad (n \geq 2) \quad \text{この漸化式を解けば } \gamma_n \text{ が計算できる。}\end{aligned}$$

これより $\gamma_0 = \frac{\phi_1^2}{1 - \phi_2} \gamma_0 + \left(\frac{\phi_1^2}{1 - \phi_2} + \phi_2 \right) \gamma_0 + \sigma^2$, すなわち、

$$\gamma_0 = \frac{\sigma^2}{1 - \frac{\phi_1^2}{1 - \phi_2} - \left(\frac{\phi_1^2}{1 - \phi_2} + \phi_2 \right)} = \frac{1 - \phi_2}{1 + \phi_2} \frac{\sigma^2}{(1 - \phi_2)^2 - \phi_1^2} > 0.$$

問題 2.1 定常 AR(2) モデル $Y_t = 3 - 0.3Y_{t-1} + 0.4Y_{t-2} + \varepsilon_t$, $E[\varepsilon_t] = 0$, $V(\varepsilon_t) = 3$ において、 $\mu = E[Y_t]$, 自己共分散 $\gamma_0, \gamma_1, \gamma_2, \gamma_n$ ($n \geq 3$), 自己相関 ρ_1, ρ_2, ρ_n ($n \geq 3$) を求めよ。

問題 2.2 定常 AR(2) モデル $Y_t = 1 + \frac{2}{3}Y_{t-1} + \frac{1}{6}Y_{t-2} + \varepsilon_t$, で、

(1) 定常性を示せ。(2) $\mu = E[Y_t]$ を求めよ。(3) $\sigma^2 = 2$ として $\gamma_0, \gamma_1, \gamma_2, \gamma_3$ を求めよ。

問題 2.3 (1) 定常 AR(2) モデル $Y_t = 1 + aY_{t-1} + 0.2Y_{t-2} + \varepsilon_t$ で、定常性をもつための a の範囲を求めよ。

(2) 定常 AR(2) モデル $Y_t = 1 - 0.5Y_{t-1} + bY_{t-2} + \varepsilon_t$ で、定常性をもつための b の範囲を求めよ。

2.3 MA(q) (q 次の移動平均モデル, Moving-average Model)

$\{Y_t\}$ が q 次の移動平均モデル MA(q) に従うとは、

$$Y_t = \theta_0 + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \cdots - \theta_q\varepsilon_{t-q} \quad (2.4)$$

となることにいう。ここで、 $\theta_0, \theta_1, \dots, \theta_q$ は定数であり、誤差項 ε_t は $Y_{t-1}, Y_{t-2}, \dots, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots$ と独立で $E[\varepsilon_t] = 0$, $V(\varepsilon_t) = \sigma^2$ とし、同一の分布に従う。 $\{Y_t\} \sim \text{MA}(q)$ と表す。

MA(q) は必ず定常性を満たす。実際、 $E[Y_t] = \theta_0$ であり、

$$\gamma_0 = V(Y_t) = (1 + \theta_1^2 + \cdots + \theta_q^2)\sigma^2,$$

$$\begin{aligned} \gamma_1 = \text{Cov}(Y_t, Y_{t-1}) &= \text{Cov}(\varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \cdots - \theta_q\varepsilon_{t-q}, \varepsilon_{t-1} - \theta_1\varepsilon_{t-2} - \theta_2\varepsilon_{t-3} - \cdots - \theta_q\varepsilon_{t-q-1}) \\ &= (-\theta_1 + \theta_1\theta_2 + \theta_2\theta_3 + \cdots + \theta_{q-1}\theta_q)\sigma^2. \end{aligned}$$

同様に $\gamma_2 = (-\theta_2 + \theta_1\theta_3 + \cdots + \theta_{q-2}\theta_q)\sigma^2, \dots, \gamma_q = -\theta_q\sigma^2, \gamma_n = 0$ ($n \geq q+1$).

自己相関 ρ_n は $\rho_0 = 1, \rho_1 = \frac{-\theta_1 + \theta_1\theta_2 + \theta_2\theta_3 + \cdots + \theta_{q-1}\theta_q}{1 + \theta_1^2 + \cdots + \theta_q^2}, \rho_2 = \frac{-\theta_2 + \theta_1\theta_3 + \cdots + \theta_{q-2}\theta_q}{1 + \theta_1^2 + \cdots + \theta_q^2}, \dots$

定常な AR(p) $\{Y_t\}$ は、 $Y_t = \xi_0 + \varepsilon_t + \sum_{i=1}^{\infty} \xi_i\varepsilon_{t-i}$ と表されることより、MA(∞).

実際、AR(p) の MA(∞) 表現を、 $Y_t = \xi_0 + \varepsilon_t + \xi_1\varepsilon_{t-1} + \xi_2\varepsilon_{t-2} + \cdots$ とすると、

$$\begin{aligned} Y_t &= \phi_0 + \phi_1Y_{t-1} + \cdots + \phi_pY_{t-p} + \varepsilon_t \\ &= \phi_0 + \phi_1(\xi_0 + \varepsilon_{t-1} + \xi_1\varepsilon_{t-2} + \xi_2\varepsilon_{t-3} + \cdots) + \phi_2(\xi_0 + \varepsilon_{t-2} + \xi_1\varepsilon_{t-3} + \xi_2\varepsilon_{t-4} + \cdots) + \\ &\quad + \cdots + \phi_p(\xi_0 + \varepsilon_{t-p} + \xi_1\varepsilon_{t-p-1} + \xi_2\varepsilon_{t-p-2} + \cdots) + \varepsilon_t \\ &= \phi_0 + (\phi_1 + \cdots + \phi_p)\xi_0 + \varepsilon_t + \phi_1\varepsilon_{t-1} + (\phi_1\xi_1 + \phi_2)\varepsilon_{t-2} + (\phi_1\xi_2 + \phi_2\xi_1 + \phi_3)\varepsilon_{t-3} + \cdots. \end{aligned}$$

よって、 $\xi_0 = \frac{\phi_0}{1 - (\phi_1 + \cdots + \phi_p)} (= \mu)$, $\xi_1 = \phi_1$, $\xi_2 = \phi_1\xi_1 + \phi_2 = \phi_1^2 + \phi_2$, $\xi_3 = \phi_1\xi_2 + \phi_2\xi_1 + \phi_3 = \cdots$.

逆に、上で Y_t と ε_t の役割を入れ替えれば、次がわかる。

命題 2.2 (反転可能性) 特性方程式 $\theta(x) = 1 - (\theta_1x + \theta_2x^2 + \cdots + \theta_qx^q) = 0$ の解の絶対値がすべて 1 より大きいとき、MA(q) は AR(∞) で表現できる。すなわち、反転可能である。

命題 2.3 (識別可能性) 特性方程式 $\theta(x) = 0$ の解の絶対値がすべて 1 以上のとき、MA(q) は識別可能である。すなわち、平均、分散、自己共分散からモデルが一意的に定まる。

例題 2.1 Y_t が MA(1), つまり $Y_t = \theta_0 + \varepsilon_t - \theta_1\varepsilon_{t-1}$ について、 $\mu = E[Y_t]$, $\gamma_0 = V(Y_t)$, $\gamma_1 = \text{Cov}(Y_t, Y_{t-1})$ のとき、 $\theta_0, \theta_1, \sigma^2$ はどうなるか。

解: 別の $\delta_0, \delta_1, \omega^2$ があつたとすると、

$$\mu = \theta_0 = \delta_0, \quad \gamma_0 = \sigma^2(1 + \theta_1^2) = \omega^2(1 + \delta_1^2), \quad \gamma_1 = (-\theta_1)\sigma^2 = (-\delta_1)\omega^2.$$

γ_0, γ_1 の式より

$$\frac{\theta_1}{1 + \theta_1^2} = \frac{\delta_1}{1 + \delta_1^2}, \quad \text{ゆえに } \theta_1(1 + \delta_1^2) - \delta_1(1 + \theta_1^2) = (\theta_1 - \delta_1)(1 - \theta_1\delta_1) = 0$$

よつて、 $\theta_1\delta_1 = 1$ であれば、別のパラメータでも同じ μ, γ_0, γ_1 をとり得る。

ここで、特性方程式 $1 - \theta_1x = 0$ の解は $x = \frac{1}{\theta_1}$ 。よつて、 $\left|\frac{1}{\theta_1}\right| \geq 1$, すなわち $|\theta_1| \leq 1$ となる方を取つてくれば、一意的に定まる。 □

問題 2.4 MA(2) モデル $Y_t = 3 + \varepsilon_t - a\varepsilon_{t-1} + a\varepsilon_{t-2}$ が反転可能性をもつ実数 a の範囲を求めよ。また、識別可能性をもつ実数 a の範囲を求めよ。

問題 2.5 MA(2) モデル $Y_t = 2 + \varepsilon_t + \frac{5}{6}\varepsilon_{t-1} - \frac{1}{6}\varepsilon_{t-2}$, $E[\varepsilon_t] = 0$, $V(\varepsilon_t) = 2$ において、 $\mu = E[Y_t]$, 自己共分散 $\gamma_0, \gamma_1, \gamma_2, \gamma_3$ を求めよ。

問題 2.6 MA(2) モデル $Y_t = 1 + \varepsilon_t - \frac{1}{4}\varepsilon_{t-1} - \frac{1}{4}\varepsilon_{t-2}$, ただし、 $\varepsilon_t \sim N(0, 3^2)$ とする。このとき、自己共分散 $\gamma_0, \gamma_1, \gamma_2, \gamma_3$ を求めよ。また、 Y_t の分布、 (Y_t, Y_{t+1}) の同時分布を求めよ。

問題 2.7 定常 AR(2) モデル $Y_t = 1 + 0.8Y_{t-1} + 0.3Y_{t-2} + \varepsilon_t$ の MA(∞) 表現 $Y_t = \xi_0 + \varepsilon_t + \xi_1\varepsilon_{t-1} + \xi_2\varepsilon_{t-2} + \dots$ における、 $\xi_0, \xi_1, \xi_2, \xi_3$ を求めよ。

問題 2.8 定常 AR(2) モデル $Y_t = 3 + \frac{5}{6}Y_{t-1} - \frac{1}{6}Y_{t-2} + \varepsilon_t$ において、次の二つの方法で Y_t の MA(∞) 表現を求めよ。

(1) $Y_t = \mu + \sum_{i=0}^{\infty} a_i\varepsilon_{t-i}$ とおくと、 $\{a_i\}$ が $a_{i+2} = \frac{5}{6}a_{i+1} - \frac{1}{6}a_i$, $a_0 = 1, a_1 = \frac{5}{6}$ を満たすことを示す。

(2) ラグ作用素 $LY_t = Y_{t-1}$ を用い、 $\phi(L)(Y_t - \mu) = \varepsilon_t$ とし、 $1/\phi(x)$ を x のべき級数で表す。

2.4 ARMA(p, q) (Autoregressive Moving-average Model)

$\{Y_t\}$ が次数 p, q の自己回帰移動平均モデル ARMA(q, p) に従うとは、

$$Y_t = c + \phi_1Y_{t-1} + \phi_2Y_{t-2} + \dots + \phi_pY_{t-p} + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \dots - \theta_q\varepsilon_{t-q} \quad (2.5)$$

となることにいう。ここで、 $c, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$ は定数であり、誤差項 ε_t は $Y_{t-1}, Y_{t-2}, \dots, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots$ と独立で $E[\varepsilon_t] = 0$, $V(\varepsilon_t) = \sigma^2$ とし、同一の分布に従う。 $\{Y_t\} \sim \text{ARMA}(q)$ と表す。

命題 2.4 (1) $\phi(x) = 1 - (\phi_1x + \phi_2x^2 + \dots + \phi_px^p) = 0$ の解の絶対値がすべて 1 より大きいとき、ARMA(p, q) は定常性をもつ。

(2) $\theta(x) = 1 - (\theta_1x + \theta_2x^2 + \dots + \theta_qx^q) = 0$ の解の絶対値がすべて 1 より大きいとき、ARMA(p, q) は反転可能である。

(3) $\theta(x) = 0$ の解の絶対値がすべて 1 以上であり、かつ $\phi(x) = 0$ と共通解をもたないとき、ARMA(p, q) は識別可能である。

例題 2.2 定常な ARMA(1, 1) モデル $Y_t = \phi_1Y_{t-1} + \varepsilon_t - \theta_1\varepsilon_{t-1}$ のパラメータ ϕ_1, θ_1 および $\sigma^2 (= V(\varepsilon_t))$ を用いて自己相関 ρ_h ($h = 1, 2, \dots$) を表せ。

解: $Y_t - \phi_1 Y_{t-1} = \varepsilon_t - \theta_1 \varepsilon_{t-1}$ で

$$\begin{aligned} V(Y_t - \phi_1 Y_{t-1}) &= V(Y_t) - 2\phi_1 \text{Cov}(Y_t, Y_{t-1}) + \phi_1^2 V(Y_{t-1}) = \gamma_0(1 + \phi_1^2) - 2\phi_1 \gamma_1, \\ V(\varepsilon_t - \theta_1 \varepsilon_{t-1}) &= \sigma^2(1 + \theta_1^2), \end{aligned}$$

より $\gamma_0(1 + \phi_1^2) - 2\phi_1 \gamma_1 = \sigma^2(1 + \theta_1^2)$. また、

$$\begin{aligned} \gamma_1 &= \text{Cov}(Y_t, Y_{t+1}) = \text{Cov}(Y_t, \phi_1 Y_t + \varepsilon_{t+1} - \theta_1 \varepsilon_t) \\ &= \phi_1 \text{Cov}(Y_t, Y_t) + \text{Cov}(\phi_1 Y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}, \varepsilon_{t+1} - \theta_1 \varepsilon_t) = \phi_1 \gamma_0 - \theta_1 \sigma^2. \end{aligned}$$

$$\begin{aligned} \text{よって、} \gamma_0(1 + \phi_1^2) - 2\phi_1(\phi_1 \gamma_0 - \theta_1 \sigma^2) &= \sigma^2(1 + \theta_1^2) \text{ より } \gamma_0 = \frac{\sigma^2(1 + \theta_1^2 - 2\phi_1 \theta_1)}{1 - \phi_1^2}, \\ \gamma_1 &= \frac{\sigma^2(\phi_1 + \phi_1 \theta_1^2 - \phi_1^2 \theta_1 - \theta_1)}{1 - \phi_1^2} = \frac{\sigma^2(\theta_1 - \phi_1)(\phi_1 \theta_1 - 1)}{1 - \phi_1^2}. \quad \therefore \rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{(\theta_1 - \phi_1)(\phi_1 \theta_1 - 1)}{1 + \theta_1^2 - 2\phi_1 \theta_1}. \end{aligned}$$

また、 $h \geq 2$ のとき、

$$\gamma_h = \text{Cov}(Y_t, Y_{t+h}) = \text{Cov}(Y_t, \phi_1 Y_{t+h-1} + \varepsilon_{t+h} - \theta_1 \varepsilon_{t+h-1}) = \phi_1 \gamma_{h-1}.$$

$$\text{よって、} \gamma_h = \phi_1^{h-1} \gamma_1 \text{ より、} \rho_h = \phi_1^{h-1} \rho_1 = \frac{(\theta_1 - \phi_1)(\phi_1 \theta_1 - 1)}{1 + \theta_1^2 - 2\phi_1 \theta_1} \quad (h \geq 2). \quad \square$$

問題 2.9 定常 ARMA(1, 1) モデル $Y_t = 2 + \frac{1}{3}Y_{t-1} - \frac{1}{2}\varepsilon_{t-1} + \varepsilon_t$, $E[\varepsilon_t] = 0$, $V(\varepsilon_t) = 4^2$ において、 $\mu = E[Y_t]$, $\gamma_h = \text{Cov}(Y_t, Y_{t-h})$ を求めよ。

2.5 時系列モデルに基づく予測

• 偏自己相関: 確率過程の自己相関を $\rho_1, \rho_2, \rho_3, \dots$ とし、 $Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t$ のとき、

$$\begin{aligned} \rho_1 &= \phi_{11} & \rho_1 &= \phi_{21} + \phi_{22} \rho_1 & \rho_1 &= \phi_{31} + \phi_{32} \rho_1 + \phi_{33} \rho_2 \\ \rho_2 &= \phi_{21} \rho_1 + \phi_{22} & \rho_2 &= \phi_{31} \rho_1 + \phi_{32} + \phi_{33} \rho_1 & \dots \\ \rho_3 &= \phi_{31} \rho_2 + \phi_{32} \rho_1 + \phi_{33} \end{aligned}$$

で定まる ϕ_{hh} を Y_t と Y_{t-h} の偏自己相関という。(ユール=ウオーカーの方程式 (2.3) で ϕ_i を ϕ_{hi} とした。)

問題 2.10 次の定常 AR(2) モデル, MA(2) モデルにおいて偏自己相関 $\phi_{11}, \phi_{22}, \phi_{33}$ を求めよ。

- (1) 定常 AR(2) モデル $Y_t = 3 - 0.3Y_{t-1} + 0.4Y_{t-2} + \varepsilon_t$, (cf. 問題 2.1).
- (2) 定常 AR(2) モデル $Y_t = 1 + \frac{2}{3}Y_{t-1} + \frac{1}{6}Y_{t-2} + \varepsilon_t$, (cf. 問題 2.2).
- (3) 定常 MA(2) モデル $Y_t = 2 + \varepsilon_t + \frac{5}{6}\varepsilon_{t-1} - \frac{1}{6}\varepsilon_{t-2}$, (cf. 問題 2.5).

• 時点 n までのデータ $\{y_t\}$ が与えられたときの h 時点先の Y_{n+h} の予測:

一般に、 Y の X による予測値として $E[Y|X]$ をとる。これを (2 乗平均) 最良予測値という。これは、 $E[(Y - g(X))^2]$ を最小にするのは $g(X) = E[Y|X]$ となることによる。また、

$$E[f(X)|X] = f(X), \quad X \text{ と } Y \text{ が独立なら } E[X|Y] = E[X]$$

に注意する。(前期に定理 1.7 として示した。)

[AR(p) の予測] $Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t$ のとき、 Y_{n+1} の予測値として次の \hat{Y}_{n+1} をとる。

$$\begin{aligned} \hat{Y}_{n+1} &= E[Y_{n+1}|Y_n, Y_{n-1}, \dots] = E[\phi_0 + \phi_1 Y_n + \phi_2 Y_{n-1} + \dots + \phi_p Y_{n-p+1} + \varepsilon_{n+1}|Y_n, Y_{n-1}, \dots] \\ &= \phi_0 + \phi_1 Y_n + \phi_2 Y_{n-1} + \dots + \phi_p Y_{n-p+1} \end{aligned}$$

最後の等号は ε_{n+1} が Y_n, Y_{n-1}, \dots と独立であることと $E[\varepsilon_{n+1}] = 0$ による。

$Y_{n+1} = \hat{Y}_{n+1} + \varepsilon_{n+1}$ に注意して

$$\begin{aligned} Y_{n+2} &= \phi_0 + \phi_1 Y_{n+1} + \phi_2 Y_n + \dots + \phi_p Y_{n-p+2} + \varepsilon_{n+2} \\ &= \phi_0 + \phi_1 (\hat{Y}_{n+1} + \varepsilon_{n+1}) + \phi_2 Y_{n-1} + \dots + \phi_p Y_{n-p+1} + \varepsilon_{n+2}, \\ \hat{Y}_{n+2} &= E[Y_{n+2} | Y_n, Y_{n-1}, \dots] = \phi_0 + \phi_1 \hat{Y}_{n+1} + \phi_2 Y_n + \dots + \phi_p Y_{n-p+2} \end{aligned}$$

同様に、 $\hat{Y}_{n+3} = \phi_0 + \phi_1 \hat{Y}_{n+2} + \phi_2 \hat{Y}_{n+1} + \phi_3 Y_n + \dots + \phi_p Y_{n-p+3}$ となる。

[MA(q) の予測量] $Y_t = \theta_0 + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$ のとき、

$$\begin{aligned} \hat{Y}_{n+1} &= E[Y_{n+1} | \varepsilon_n, \varepsilon_{n-1}, \dots] = \theta_0 - \theta_1 \varepsilon_n - \dots - \theta_{q-1} \varepsilon_{n+2-q} - \theta_q \varepsilon_{n+1-q}, \\ \hat{Y}_{n+2} &= E[Y_{n+2} | \varepsilon_n, \varepsilon_{n-1}, \dots] = \theta_0 - \theta_2 \varepsilon_n - \dots - \theta_q \varepsilon_{n+2-q}, \\ &\vdots \\ \hat{Y}_{n+q} &= E[Y_{n+2} | \varepsilon_n, \varepsilon_{n-1}, \dots] = \theta_0 - \theta_q \varepsilon_n, \\ \hat{Y}_{n+h} &= 0, \quad h = q+1, q+2, \dots \end{aligned}$$

すなわち、 $\varepsilon_{n+1}, \dots, \varepsilon_{n+h}$ をゼロとしたものになる。

なお、誤差項 $\varepsilon_n, \varepsilon_{n-1}, \dots$ は直接観察されるものではないため、パラメータ推定等の結果として得られる残差 e_n, e_{n-1}, \dots を用いて計算するものとなる。(残差についてはモデリングの教科書 p.2-22 を参照のこと。)

問題 2.11 定常 AR(1) モデル $Y_t = 2 + \frac{1}{3} Y_{t-1} + \varepsilon_t$, $\varepsilon_t \sim N\left(0, \frac{1}{5}\right)$, について、 $Y_T = 2.7$ が与えられたとする。

- (1) 最良予測値 \hat{Y}_{T+1} を求め、2乗平均誤差 $E[(Y_{T+1} - \hat{Y}_{T+1})^2 | Y_T = 2.7]$ を求めよ。
また、最良予測値 \hat{Y}_{T+2} を求め、2乗平均誤差 $E[(Y_{T+2} - \hat{Y}_{T+2})^2 | Y_T = 2.7]$ を求めよ。
(2) Y_{T+1} と Y_{T+2} の信頼度 95% の信頼区間を求めよ。ただし小数点以下第 2 位まで求めよ。

問題 2.12 定常 MA(2) モデル $Y_t = 1 + \varepsilon_t - \frac{2}{3} \varepsilon_{t-1} - \frac{1}{3} \varepsilon_{t-2}$, $\varepsilon_t \sim N\left(0, \frac{1}{4}\right)$, について、 $\varepsilon_{T-2} = -\frac{1}{3}, \varepsilon_{T-1} = \frac{1}{3}, \varepsilon_T = \frac{2}{3}$ が与えられたとする。

- (1) 最良予測値 \hat{Y}_{T+1} を求め、2乗平均誤差 $E[(Y_{T+1} - \hat{Y}_{T+1})^2 | \varepsilon_{T-2} = -\frac{1}{3}, \varepsilon_{T-1} = \frac{1}{3}, \varepsilon_T = \frac{2}{3}]$ を求めよ。
また、最良予測値 \hat{Y}_{T+2} を求め、2乗平均誤差 $E[(Y_{T+2} - \hat{Y}_{T+2})^2 | \varepsilon_{T-2} = -\frac{1}{3}, \varepsilon_{T-1} = \frac{1}{3}, \varepsilon_T = \frac{2}{3}]$ を求めよ。
(2) Y_{T+1} と Y_{T+2} の信頼度 95% の信頼区間を求めよ。ただし小数点以下第 3 位まで求めよ。

2.6 統計の復習 2 順序統計量

前期に扱った 1.7 順序統計量を復習する。

X_1, \dots, X_n が i.i.d. で連続型とする。 $X_{(1)}, \dots, X_{(n)}$ を小さいものから並べたものを

$$X_{(1)}, X_{(2)}, \dots, X_{(n)}$$

とし、 $X_{(j)}$ を j 番目の順序統計量という。 X_1 の (共通の) 密度関数を $f(x)$, 分布関数を $F(x)$ とする。

明らかに、 $X_{(1)} = \min\{X_1, \dots, X_n\}$, $X_{(n)} = \max\{X_1, \dots, X_n\}$ である。

また、 $n = 2m - 1$ (奇数) のとき $X_{(m)}$, $n = 2m$ (偶数) のとき $\frac{1}{2}\{X_{(m)} + X_{(m+1)}\}$ が X_1, \dots, X_n の中央値であった。

対称性に注意すると、 $(X_{(1)}, \dots, X_{(n)})$ の同時密度関数 g が

$$g(t_1, t_2, \dots, t_n) = \begin{cases} n! f(t_1) f(t_2) \dots f(t_n) & (t_1 < t_2 < \dots < t_n) \\ 0 & (\text{その他}) \end{cases}$$

となることがわかる。また、同様に対称性より $x < y$ のとき、

$$\iint \cdots \int_{x < t_1 < t_2 < \cdots < t_k < y} f(t_1)f(t_2)\cdots f(t_k) dt_1 dt_2 \cdots dt_k = \frac{1}{k!} \{F(y) - F(x)\}^k$$

ここで、 $F(x)$ は X_1 の分布関数である。よって、 $X_{(j)}$ の密度関数は、この周辺密度なので

$$\begin{aligned} f_{X_{(j)}}(x) &= n! \int \cdots \int_{t_1 < \cdots < t_{j-1} < x} f(t_1) \cdots f(t_{j-1}) dt_1 \cdots dt_{j-1} f(x) \int \cdots \int_{x < t_{j+1} < \cdots < t_n} f(t_{j+1}) \cdots f(t_n) dt_{j+1} \cdots dt_n \\ &= \frac{n!}{(j-1)!1!(n-j)!} F(x)^{j-1} f(x) (1-F(x))^{n-j} \end{aligned}$$

となる。また、 $i < j$ とすると $(X_{(i)}, X_{(j)})$ の同時密度関数は $x < y$ に対しては

$$f_{(X_{(i)}, X_{(j)})}(x, y) = \frac{n!}{(i-1)!1!(j-1-i)!1!(n-j)!} F(x)^{i-1} f(x) (F(y) - F(x))^{j-1-i} f(y) (1-F(y))^{n-j},$$

$x \geq y$ に対しては $f_{(X_{(i)}, X_{(j)})}(x, y) = 0$ となる。

独立な $U_1, \dots, U_n \sim U(0, 1)$ の順序統計量 $U_{(k)}$ がベータ分布 $\text{BETA}(k, n-k+1)$ に従うこと等は前期に示した (cf. 前期の講義ノート pp.20-21)。また、これは 二項母集団の母比率の区間推定 (精密法) に応用された (cf. 前期の講義ノート pp.27-28, 検定は p.36)。

問題 2.13 ある人の射撃命中率は 70% であるといわれている。この人のある日の成績は 6 発中 2 発が命中したのみであった。この日はとくに命中率が低かったといえるか。5% の有意水準で検定せよ。

ここでは、指数分布に従う母集団について考察する。

指数分布に従う母集団の母平均の区間推定

- 指数分布 $\text{Ex}(\lambda)$, ガンマ分布 $\Gamma(a, \lambda)$, カイ二乗分布 χ_n^2 , ベータ分布 $\text{BETA}(a, b)$ の復習

$$X \sim \Gamma(a, \lambda) \text{ とは } X \text{ の密度関数が } f_X(x) = \begin{cases} \frac{\lambda^a}{\Gamma(a)} x^{a-1} e^{-\lambda x} & (x > 0) \\ 0 & (x \leq 0) \end{cases} \text{ とする。}$$

(a) 指数分布 $\text{Ex}(\lambda)$ とガンマ分布 $\Gamma(1, \lambda)$ は同じ分布、カイ二乗分布 χ_n^2 とガンマ分布 $\Gamma(\frac{n}{2}, \frac{1}{2})$ は同じ分布。

(b) X, Y が独立で、 $X \sim \Gamma(a, \lambda)$, $Y \sim \Gamma(b, \lambda)$ とし、 $c > 0$ を定数とすると次が成立:

$$(i) cX \sim \Gamma(a, \lambda/c), \quad (ii) X + Y \sim \Gamma(a + b, \lambda), \quad (iii) \frac{X}{X + Y} \sim \text{BETA}(a, b),$$

$$(iv) a = m/2, b = n/2 \text{ なら } W = \frac{Y/n}{X/m} \sim F_m^n \text{ であるから, } \frac{X}{X + Y} = \frac{1}{1 + Y/X} = \frac{1}{1 + \frac{n}{m}W} \text{ と変形すれば, F 分布 } F_m^n \text{ とベータ分布 } \text{BETA}(\frac{m}{2}, \frac{n}{2}) \text{ の関係も従う。}$$

命題 2.5 X_1, \dots, X_n は i.i.d. で $\text{Ex}(1/\lambda)$ に従うとき、 $\frac{2}{\lambda}(X_1 + \cdots + X_n) \sim \chi_n^2$ 。

証明は上記 (a) の前半と、(b) の (i), (ii) より容易。

この命題より $S_n = X_1 + \cdots + X_n$ とすると、

$$P\left(\chi_{2n}^2(1 - \frac{1}{2}\varepsilon) \leq \frac{2}{\lambda} S_n \leq \chi_{2n}^2(\frac{1}{2}\varepsilon)\right) = 1 - \varepsilon$$

となるから、括弧内の不等式を λ について解き、 $S_n = X_1 + \cdots + X_n$ の実現値 s に置き換えることで、平均値 λ の信頼係数 $1 - \varepsilon$ の信頼区間が $\frac{2s}{\chi_{2n}^2(\frac{1}{2}\varepsilon)} \leq \lambda \leq \frac{2s}{\chi_{2n}^2(1 - \frac{1}{2}\varepsilon)}$ となることがわかる。

また、検定も同様にできる。帰無仮説 $H_0: \lambda = \lambda_0$, 対立仮説 $H_1: \lambda \neq \lambda_0$ を有意水準 ε で検定するには、実現値を $s = x_1 + \cdots + x_n$ とするとき、

$$\frac{2}{\lambda_0} s < \chi_{2n}^2(1 - \varepsilon/2) \text{ または } \chi_{2n}^2(\varepsilon/2) < \frac{2}{\lambda_0} s \text{ のとき } H_0 \text{ を棄却し,} \quad (2.6)$$

$$\chi_{2n}^2(1-\varepsilon/2) < \frac{2}{\lambda_0} s < \chi_{2n}^2(\varepsilon/2) \text{ のとき } H_0 \text{ を採択する。} \quad (2.7)$$

片側検定の場合も同様に考えることができる。(各自考えよ。)

次に、データが n 個で打ち切られた場合 を考える。

命題 2.6 X_1, \dots, X_N は i.i.d. で $\text{Ex}(1/\lambda)$ に従うとし、 j 番目の順序統計量を $T_j = X_{(j)}$, $j = 1, \dots, n (< N)$, とする。

(1) (T_1, T_2, \dots, T_n) の密度関数 $f(t_1, t_2, \dots, t_n)$ は

$$f(t_1, \dots, t_n) = \begin{cases} \frac{N!}{(N-n)! \lambda^n} \exp\left[-\frac{1}{\lambda} \left\{ \sum_{j=1}^n t_j + (N-n)t_n \right\}\right] & (0 < t_1 < t_2 < \dots < t_n) \\ 0 & (\text{その他}) \end{cases} \quad (2.8)$$

(2) (2.8) を尤度関数とする λ の最尤推定量は $\hat{\lambda} = \frac{1}{n} \left\{ \sum_{j=1}^n T_j + (N-n)T_n \right\}$.

(3) (2) の $\hat{\lambda}$ に対し、 $2n \cdot \frac{\hat{\lambda}}{\lambda} \sim \chi_{2n}^2$.

略証: (1) $P(T_k \geq t_n) = e^{-\frac{1}{\lambda} t_n}$ と順序統計量の定義から容易に示せる。

(2) 対数尤度関数 $l(\lambda) = \log f(t_1, \dots, t_n)$ を λ に関し微分せよ。

(3) $s < 1/2$ に対し $E[e^{s(2n\hat{\lambda}/\lambda)}] = (1-2s)^{-n}$ (χ_{2n}^2 分布の積率母関数) を示せばよい。

$$\begin{aligned} E[e^{s(2n\hat{\lambda}/\lambda)}] &= E[e^{\frac{2s}{\lambda} \{ \sum_{j=1}^n T_j + (N-n)T_n \}}] \\ &= \frac{N!}{(N-n)! \lambda^n} \int_0^\infty e^{-\frac{1-2s}{\lambda} t_1} dt_1 \int_{t_1}^\infty e^{-\frac{1-2s}{\lambda} t_2} dt_2 \cdots \int_{t_{n-2}}^\infty e^{-\frac{1-2s}{\lambda} t_{n-1}} dt_{n-1} \int_{t_{n-1}}^\infty e^{-\frac{1-2s}{\lambda} (N-n+1)t_n} dt_n \\ &= \cdots = (1-2s)^{-n}. \quad (n=4 < N \text{ として各自証明せよ。}) \quad \square \end{aligned}$$

この命題より $S = T_1 + \cdots + T_n + (N-n)T_n$ とすると、

$$P\left(\chi_{2n}^2(1-\frac{1}{2}\varepsilon) \leq \frac{2}{\lambda} S \leq \chi_{2n}^2(\frac{1}{2}\varepsilon)\right) = 1-\varepsilon.$$

よって、 $S = T_1 + \cdots + T_n + (N-n)T_n$ の実現値をとすると、平均値 λ の信頼係数 $1-\varepsilon$ の信頼区間が

$$\frac{2s}{\chi_{2n}^2(\frac{1}{2}\varepsilon)} \leq \lambda \leq \frac{2s}{\chi_{2n}^2(1-\frac{1}{2}\varepsilon)} \quad (2.9)$$

となることがわかる。

検定についても同様にできる。帰無仮説 $H_0: \lambda = \lambda_0$, 対立仮説 $H_1: \lambda \neq \lambda_0$ を有意水準 ε で検定するには、実現値 $s = t_1 + \cdots + t_n + (N-n)t_n$ に対し、(2.6), (2.7) とすればよい。

データが X で打ち切られた場合 (寿命試験では定時打ち切りという)

大きさ N の標本がその値の X 以下であるもののみが測定され、観測された n 個の値は x_1, \dots, x_n であったとする。このときは $s = x_1 + \cdots + x_n + (N-n)X$ をその実現値として、

平均値 λ の信頼係数 $1-\varepsilon$ の信頼区間は (2.9) となる。

帰無仮説 $H_0: \lambda = \lambda_0$, 対立仮説 $H_1: \lambda \neq \lambda_0$ を有意水準 ε で検定するには、(2.6), (2.7) とすればよい。

問題 2.14 以下の問いに答えよ。(1), (2) は小数点以下第 1 位まで求めよ。

(1) ある電気商品の寿命を調べたところ、次の通りであった。平均寿命の 95% 信頼区間を求めよ。

(*) 732, 838, 915, 1211, 1355, 1420, 1638 (単位 時間)

(2) 10 個の商品について調べた結果、(1) の (*) のような標本が得られた。3 個の部品はいずれも 1638 時間以上もったものとすれば、平均寿命の 95% 信頼区間を求めよ。

(3) 10 個の商品について調べた結果、(1) の (*) のような標本が得られ、残りの 3 個の部品はいずれも 1700 時間以上もったものとする。この商品の平均寿命は 1400 時間とってよいか。有意水準 5% で検定せよ。

3 確率過程

3.1 マルコフ連鎖とマルチンゲール

定義 3.1 $\{X_t\}$ がマルコフ過程 (連鎖) であるとは、任意の $0 \leq t_1 < t_2 < \dots < t_n < t$ と x_1, x_2, \dots, x_n, x に対して、

$$P(X_t \leq x | X_{t_1} = x_1, X_{t_2} = x_2, \dots, X_{t_n} = x_n) = P(X_t \leq x | X_{t_n} = x_n)$$

となることをいう。つまり、将来は現在のみに依存し、過去には依存しない確率過程をいう。

このとき、 $\{X_t\}$ は次のようになる:

$$\begin{aligned} & P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) \\ &= P(X_n = x_n | X_1 = x_1, \dots, X_{n-1} = x_{n-1}) P(X_1 = x_1, \dots, X_{n-1} = x_{n-1}) \\ &= P(X_n = x_n | X_{n-1} = x_{n-1}) P(X_1 = x_1, \dots, X_{n-1} = x_{n-1}) \\ &\quad \vdots \\ &= P(X_n = x_n | X_{n-1} = x_{n-1}) \cdots P(X_3 = x_3 | X_2 = x_2) P(X_2 = x_2 | X_1 = x_1). \end{aligned}$$

推移確率 $P(X_{t+1} = y | X_t = x) = p_{xy}$, 推移確率密度関数 $f_{X_{t+s}|X_t}(y|x) = \frac{f_{(X_t, X_{t+s})}(x, y)}{f_{X_t}(x)} = p(s, x, y)$ が t によらないマルコフ過程は時間的一様 (推移確率が定常) なマルコフ過程という。ここではこの場合のみを扱う。

- 有限マルコフ連鎖 X_t が $S = \{1, 2, \dots, m\}$ に値をとるの場合を考える。(S を状態空間という。)

$p_{ij} = P(X_{t+1} = j | X_t = i)$ とし、 $P = \begin{pmatrix} p_{11} & \cdots & p_{1m} \\ \vdots & & \vdots \\ p_{m1} & \cdots & p_{mm} \end{pmatrix}$ とおく。 P を推移確率行列という。

${}_n p_{ij} = P(X_{t+n} = j | X_t = i)$ とすると、 $P^n = \begin{pmatrix} {}_n p_{11} & \cdots & {}_n p_{1m} \\ \vdots & & \vdots \\ {}_n p_{m1} & \cdots & {}_n p_{mm} \end{pmatrix}$ となる。特に、チャップマン-コロモゴ

ロフの方程式 ${}_n p_{ij} = \sum_{k=1}^m {}_{n-1} p_{ik} {}_1 p_{kj}$ が成り立つ。これは

$$P(X_{t+n_1+n_2} = j | X_t = i) = \sum_{k=1}^m P(X_{t+n_1} = k | X_t = i) P(X_{t+n_1+n_2} = j | X_{t+n_1} = k)$$

から容易にわかる。また、 $P(X_0 = i) = p_i$ とすると、

$$(P(X_n = 1), \dots, P(X_n = m)) = (P(X_{n-1} = 1), \dots, P(X_{n-1} = m)) P = (p_1, \dots, p_m) P^n.$$

特に、 $\pi = (\pi_1, \dots, \pi_m)$ が $\pi = \pi P$ を満たすとき、 π を定常分布という。有限マルコフ連鎖では、 $\forall i, j$ に対してある n があって ${}_n p_{ij} > 0$ とできる (このときマルコフ連鎖は既約であるという) とき定常分布はただ一つ

であり、特にある n_0 があって $\forall i, j$ に対し ${}_{n_0} p_{ij} > 0$ を満たせば $\lim_{n \rightarrow \infty} P^n = \begin{pmatrix} \pi_1 & \cdots & \pi_m \\ \vdots & & \vdots \\ \pi_1 & \cdots & \pi_m \end{pmatrix}$ となる。

- 2重マルコフ過程

$$P(X_{t+1} \leq x | X_0 = x_0, X_1 = x_1, \dots, X_{t-1} = x_{t-1}, X_t = x_t) = P(X_{t+1} \leq x | X_t = x_t, X_{t-1} = x_{t-1})$$

が成立しているとき、 X_t を 2重マルコフ過程という。この場合は $Y_t = (X_t, X_{t+1})$ とペアにして考えると Y_t はマルコフ過程となる。

例題 3.1 X_t を $\{1, 2, 3\}$ 上のマルコフ連鎖でその推移確率を $P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \end{pmatrix}$ とする。

$P(X_n = i | X_0 = j) = {}_n p_{ji}$ とするとき、(1) ${}_2 p_{11}, {}_3 p_{11}$ を求めよ。(2) ${}_n p_{11}$ を求めよ。

解: (1) ${}_2 p_{11} = p_{11}p_{11} + p_{12}p_{21} + p_{13}p_{31} = 0 \cdot 0 + 1 \cdot 0 + 0 \cdot \frac{1}{2} = 0,$

また、 ${}_2 p_{21} = 0 \cdot 0 + \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}, {}_2 p_{31} = \frac{1}{2} \cdot 0 + 0 \cdot 0 + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$ より、

${}_3 p_{11} = p_{11} \cdot {}_2 p_{11} + p_{12} \cdot {}_2 p_{21} + p_{13} \cdot {}_2 p_{31} = 0 \cdot 0 + 1 \cdot \frac{1}{4} + 0 \cdot \frac{1}{4} = \frac{1}{4}.$

(2) P の固有値は $1, i/2, -i/2$. よって、すべての固有値で重複度 1 より P は対角化可能なので、正則行列 U があって、 $P = UDU^{-1}$ とできる*2。ただし、 D は $1, i/2, -i/2$ を対角成分とする対角行列である。よって、 $P^n = U D^n U^{-1}$ で D^n は $1, (i/2)^n, (-i/2)^n$ を対角成分とする対角行列なので、定数 a, b, c があって、

$${}_n p_{11} = a + b \left(\frac{i}{2}\right)^n + c \left(-\frac{i}{2}\right)^n$$

と表せる。ここでは、 ${}_n p_{11}$ は実数で、 $\pm i = e^{\pm i\pi/2}$ より

$$\left(\pm \frac{i}{2}\right)^n = \left(\frac{1}{2}\right)^n \left(e^{\pm i\pi/2}\right)^n = \left(\frac{1}{2}\right)^n \left(\cos \frac{n\pi}{2} \pm i \sin \frac{n\pi}{2}\right)$$

となるので、定数 α, β, γ があって、

$${}_n p_{11} = \alpha + \left(\frac{1}{2}\right)^n \left\{ \beta \cos \frac{n\pi}{2} + \gamma \sin \frac{n\pi}{2} \right\}. \quad (3.1)$$

と表せる。ここで、 $P^0 = ({}_0 p_{ij})$ は単位行列、 $P = ({}_1 p_{ij})$ に注意する。よって、

$$1 = {}_0 p_{11} = \alpha + \beta, \quad 0 = {}_1 p_{11} = \alpha + \frac{1}{2}\gamma, \quad 0 = {}_2 p_{11} = \alpha - \frac{1}{4}\beta$$

となるので $\alpha = 1/5, \beta = 4/5, \gamma = -2/5$. よって、

$$P_1(X_n = 1) = p_{11}^{(n)} = \frac{1}{5} + \left(\frac{1}{2}\right)^n \left\{ \frac{4}{5} \cos \frac{n\pi}{2} - \frac{2}{5} \sin \frac{n\pi}{2} \right\}. \quad \square$$

問題 3.1 例 3.1 のマルコフ連鎖 X_t に対して (1) ${}_2 p_{22}, {}_2 p_{23}, {}_3 p_{23}$ を求めよ。(2) ${}_n p_{23}$ を求めよ。

ヒント: ${}_n p_{23}$ も (3.1) の右辺のように表せる。

例題 3.2 次の推移行列をもつ $\{1, 2, 3, 4\}$ 上のマルコフ連鎖の定常分布を求めよ。(b), (c) は問題 3.2 とする。)

$$(a) \begin{pmatrix} 0.4 & 0.6 & 0 & 0 \\ 0 & 0.7 & 0.3 & 0 \\ 0.1 & 0 & 0.4 & 0.5 \\ 0.5 & 0 & 0 & 0.5 \end{pmatrix} \quad (b) \begin{pmatrix} 0.7 & 0.3 & 0 & 0 \\ 0.2 & 0.5 & 0.3 & 0 \\ 0 & 0.3 & 0.6 & 0.1 \\ 0 & 0 & 0.2 & 0.8 \end{pmatrix} \quad (c) \begin{pmatrix} 0.7 & 0 & 0.3 & 0 \\ 0.2 & 0.5 & 0.3 & 0 \\ 0.1 & 0.2 & 0.4 & 0.3 \\ 0 & 0.4 & 0 & 0.6 \end{pmatrix}$$

解: 定常分布を $\pi = (\pi_1, \pi_2, \pi_3, \pi_4)$ とする。 $\pi P = \pi$ を書き直して、

$$0.4\pi_1 + 0.1\pi_3 + 0.5\pi_4 = \pi_1, \quad 0.6\pi_1 + 0.7\pi_2 = \pi_2, \quad 0.3\pi_2 + 0.4\pi_3 = \pi_3, \quad 0.5\pi_3 + 0.5\pi_4 = \pi_4,$$

最初の 3 つの式より $\pi_1 = \pi_2/2 = \pi_3 = \pi_4$ を得る。従って、 $\pi_1 + \pi_2 + \pi_3 + \pi_4 = 1$ より、 $\pi = (1/5, 2/5, 1/5, 1/5)$ となる。 \square

問題 3.2 例 3.2 の (b), (c) を推移行列とするマルコフ連鎖の定常分布をそれぞれ求めよ。

*2 ここでは対角化やジョルダン標準形における正則行列 U の存在だけ用い、その具体形を必要としない計算法を紹介する。

定義 3.2 $\{X_t\}$ がマルチンゲールであるとは、すべての t について

$$E[X_{t+1}|X_t, X_{t-1}, \dots, X_2, X_1] = X_t, \quad \text{a.s.} \quad (3.2)$$

となることをいう。

マルチンゲールを扱うのに次の条件つき期待値についての定理が有用である (cf. 前期の**定理 1.7**)。

定理 3.1 条件つき期待値について以下が成立する。ただし、 X は多次元確率変数であってもよい。

- (1) $E[E[Y|X]] = E[Y]$,
- (2) $E[aY + bZ|X] = aE[Y|X] + bE[Z|X]$, a.s. (a, b は定数。)
- (3) $E[g(X)h(Y)|X] = g(X)E[h(Y)|X]$, a.s. 特に $E[g(X)|X] = g(X)$, a.s.
- (4) X, Y が独立なら $E[Y|X] = E[Y]$, a.s.
- (5) $E[E[Z|X, Y]|X] = E[Z|X]$, a.s.

$\{X_t\}$ がマルチンゲールするとき、 $t > s$ に対して $E[X_t|X_s, X_{s-1}, \dots, X_1] = X_s$ となる。実際、

$$\begin{aligned} E[X_{s+2}|X_s, X_{s-1}, \dots, X_1] &= E[E[X_{s+2}|X_{s+1}, X_s, X_{s-1}, \dots, X_1]|X_s, X_{s-1}, \dots, X_1] \\ &= E[X_{s+1}|X_s, X_{s-1}, \dots, X_1] = X_s \end{aligned}$$

を繰り返して証明される。ここで、最初の等号では定理 3.1 (5) を次の等号は (3.2) を $t = s + 1$ とし次の等号は $t = s$ とし用いた。また、その両辺の期待値をとれば、定理 3.1(1) より $E[X_t] = E[X_s]$ が従う。

マルチンゲールの意味: 時刻 s までの情報のもとで未来 ($t > s$) の期待財産と時刻 s での財産が一致する、すなわち、 s から t までの条件付き期待財産増分が 0 であるときをいう。たとえば、公平な賭けを行っているギャンブラーの財産過程がその典型例である。

● 1次元対称ランダムウォーク

a を整数、 ξ_1, ξ_2, \dots , を i.i.d. で $P(\xi_i = 1) = P(\xi_i = -1) = \frac{1}{2}$ とする。このとき、時間パラメータ t を 0 以上の整数として、 $Z_t^a = a + \xi_1 + \xi_2 + \dots + \xi_t$ を a を出発する 1次元対称ランダムウォークという。特に $a = 0$ のとき、 $Z_t = Z_t^0$ と表し、(1次元) 対称ランダムウォークという。

Z_t は \mathbf{Z} 上のマルコフ連鎖で、その推移確率は $P(Z_{t+1} = y|Z_t = x) = \begin{cases} \frac{1}{2} & y = x, x + 1 \\ 0 & \text{その他} \end{cases}$ となる。また、

$$\begin{aligned} E[Z_{t+1}|Z_t, Z_{t-1}, \dots, Z_1] &= E[Z_t|Z_t, Z_{t-1}, \dots, Z_1] + E[\xi_{t+1}|Z_t, Z_{t-1}, \dots, Z_1] \\ &= Z_t + E[\xi_{t+1}] = Z_t \end{aligned}$$

より対称ランダムウォークはマルチンゲールとなる。ここで、最初の等号は定理 3.1(1) を次の等号ではそれぞれ項に対して (3), (4) を用いた。

問題 3.3 Z_t を対称ランダムウォークとするとき、次を求めよ。($0 < s < t$, a, b を自然数とする。)

- (1) $E[Z_t - Z_s]$, (2) $V(Z_t - Z_s)$, (3) $P(Z_{a+b} = a - b)$, (4) $P(Z_5 = 3, Z_8 = 2)$

ヒント: (4) は Z_t の独立増分性を用いよ。

問題 3.4 ξ_1, ξ_2, \dots , を i.i.d. で $P(\xi_i = 1) = P(\xi_i = -1) = \frac{1}{2}$ とし、 $Z_t = \xi_1 + \xi_2 + \dots + \xi_t$ とするとき、

$$E[Z_{t+1}^2 - (t+1)|\xi_t, \xi_{t-1}, \dots, \xi_1] = Z_t^2 - t.$$

を示せ。(これを $Z_t^2 - t$ が $\xi_t, \xi_{t-1}, \dots, \xi_1$ に関してマルチンゲールであるという。)

「藤田岳彦 著 確率・統計・モデリング問題集」 p.36- の練習問題 3.1-9, 3.20-も上記の知識で解けます。特に、練習問題 3.25-は漸化式に関する問題なので是非挑戦ください。

3.2 ポアソン過程

定義 3.3 $N_t, t \geq 0$ がパラメータ (強度) λ のポアソン過程であるとは、

- (1) $N_0 = 0$ で $N_t \sim Po(\lambda t)$ ($t \geq 0$), つまり $P(N_t = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, k = 0, 1, 2, \dots$
- (2) (独立増分性) $0 < t_1 < t_2 < \dots < t_n$ として、 $N_{t_1}, N_{t_2} - N_{t_1}, \dots, N_{t_n} - N_{t_{n-1}}$ は独立。
- (3) (定常増分性) $t > s > 0$ として、 $N_t - N_s \sim N_{t-s}$ 。

が成立するときをいう。

モデリングの教科書 p.3-7 にあるように (1) の代わりに、

(1') (微笑時間でジャンプの起こる確率は小さい) $P(N_{t+h} - N_t \geq 2) = o(h)$, すなわち、 $\lim_{h \rightarrow 0} \frac{N_{t+h} - N_t}{h} = 0$ を仮定してもよい。((1'), (2), (3) から (1) の導出は「藤田岳彦 著 確率・統計・モデリング問題集」p.33-を参照のこと。)

例題 3.3 N_t をパラメータ λ のポアソン過程とする。次を求めよ。ただし、 $0 < s < t, k, l = 0, 1, 2, \dots$ とする。

- (1) $P(N_s = k, N_t = k + l)$, (2) $E[N_s N_t]$, (3) $\text{Cov}(N_s, N_t)$ 。

解: N_s と $N_t - N_s$ が独立であることと $N_s \sim Po(\lambda s), N_t - N_s \sim N_{t-s} \sim Po(\lambda(t-s))$ を用いる。

$$(1) P(N_s = k, N_t = k + l) = P(N_s = k, N_t - N_s = l) = P(N_s = k)P(N_t - N_s = l) = \frac{(\lambda s)^k}{k!} \frac{(\lambda(t-s))^l}{l!} e^{-\lambda t}.$$

(2) $N \sim Po(\lambda)$ のとき $E[N] = V(N) = \lambda$ に注意する。

$$\begin{aligned} E[N_t N_s] &= E[(N_t - N_s)N_s] + E[N_s^2] = E[N_t - N_s]E[N_s] + V(N_s) + (E[N_s])^2 \\ &= \lambda(t-s)\lambda s + \lambda s + (\lambda s)^2 = \lambda^2 s t + \lambda s. \end{aligned}$$

$$(3) \text{Cov}(N_s, N_t) = E[N_s N_t] - E[N_s]E[N_t] = \lambda^2 s t + \lambda s - \lambda s \lambda t = \lambda s. \quad \square$$

例 3.4 N_t をパラメータ λ のポアソン過程とすると、 $N_t - \lambda t$ はマルチンゲールとなる。

証明: $s < t$ とする。独立増分性より、 $N_t - N_s$ と $N_u, 0 \leq u \leq s$ は独立。よって、

$$\begin{aligned} E[N_t - \lambda t | N_u, 0 \leq u \leq s] &= E[N_t - N_s + N_s - \lambda t | N_u, 0 \leq u \leq s] \\ &= E[N_t - N_s] + N_s - \lambda t = \lambda(t-s) + N_s - \lambda t = N_s - \lambda s \quad \square \end{aligned}$$

問題 3.5 N_t をパラメータ λ のポアソン過程、 $0 < s < t$ とするとき次を求めよ。

- (a) $E[N_t(N_t + 1)]$, (b) $E[N_s^2 N_t]$, (c) $E[N_s N_t^2]$, (d) $E[N_t | N_s = 1]$, (e) $E[N_s | N_t = 2]$

● 計数過程としてのポアソン過程

$T_i = i$ 番目のイベントが発生する時刻とし、 n_t を時刻 t までに起こったイベントの発生回数とする。このとき n_t を計数過程 (counting process) といい、特にそのイベントの発生間隔 $T_1, T_2 - T_1, T_3 - T_2, \dots$ が独立で同分布のときの n_t を再生過程という。

ポアソン過程は発生の間隔 $T_1, T_2 - T_1, T_3 - T_2, \dots$ が独立で指数分布 $Ex(\lambda)$ の場合である。実際、 $n_t \leq k$ とは時刻 t までに $k + 1$ 回目のイベントが起こっていない、すなわち、 $T_{k+1} > t$ となるので、 $T_n \sim \Gamma(n, \lambda)$ より、

$$P(n_t \leq k) = P(T_{k+1} > t) = \int_t^\infty \frac{\lambda^{k+1}}{k!} x^k e^{-\lambda x} dx = (\text{部分積分を繰り返す}) = \sum_{j=0}^k \frac{(\lambda t)^j}{j!} e^{-\lambda t}. \quad (3.3)$$

問題 3.6 $T_1, T_2 - T_1, T_3 - T_2$ が独立で指数分布 $Ex(\lambda)$ に従うとき、 $P(3T_1 < T_2), P(3T_1 < 2T_2 < T_3), P(2T_2 < T_3)$ を求めよ。

Poisson 分布とカイ二乗分布の関係式

$T_{k+1} \sim \Gamma(k+1, \lambda)$ のとき $Y_2 = 2\lambda T_{k+1} \sim \chi_{2(k+1)}^2$ より、 $X \sim Po(\lambda)$ のとき (3.3) で $t = 1$ とすると

$$P(X \leq k) = P(T_{k+1} > 1) = P(Y_2 > 2\lambda)$$

となる。また、

$$P(X \geq k) = 1 - P(X \leq k-1) = 1 - P(Y_1 > 2\lambda) = P(Y_1 \leq 2\lambda),$$

ここで、 $Y_2 \sim \chi_{2k}^2$ となる。これより、Poisson 分布に従う母集団の母平均の区間推定 (精密法) の式、 $\frac{1}{2n}\chi_{2s}^2(1-\frac{1}{2}\varepsilon) \leq \lambda \leq \frac{1}{2n}\chi_{2(s+1)}^2(\frac{1}{2}\varepsilon)$ を得る。ここで、 s は標本を x_1, \dots, x_n , その総和が $s = \sum_{k=1}^n x_k$ を表す。

検定は、帰無仮説 $H_0: \lambda = \lambda_0$, 対立仮説 $H_1: \lambda \neq \lambda_0$ を有意水準 ε で検定するには、

$2n\lambda_0 < \chi_{2s}^2(1-\varepsilon/2)$ または $2n\lambda_0 > \chi_{2(s+1)}^2(\varepsilon/2)$ のとき H_0 を棄却する。

$\chi_{2s}^2(1-\varepsilon/2) < 2n\lambda_0 < \chi_{2(s+1)}^2(\varepsilon/2)$ のとき H_0 を採択する。

3.3 ブラウン運動

$W_t, t \geq 0$ が (標準) ブラウン運動であるとは、

(1) $W_0 = 0$ で $W_t \sim N(0, t)$ ($t \geq 0$).

(2) (独立増分性) $0 < t_1 < t_2 < \dots < t_n$ として、 $W_{t_1}, W_{t_2} - W_{t_1}, \dots, W_{t_n} - W_{t_{n-1}}$ は独立。

(3) (定常増分性) $t > s > 0$ として、 $W_t - W_s \sim W_{t-s}$.

となる時にいう。このとき、 $t \mapsto W_t$ が (確率 1 で) 連続となることが知られている。

例題 3.5 W_t を標準ブラウン運動とする。次を求めよ。ただし、 $0 < s < t$ とする。

- (1) W_t の密度関数 $f_{W_t}(x)$, (2) (W_s, W_t) の同時密度関数 $f_{(W_s, W_t)}(x, y)$, (3) $E[W_s W_t]$,
 (4) $E[W_t | W_u, 0 \leq u \leq s] = W_s$, すなわち、 W_t はマルチンゲールになる。

解: W_s と $W_t - W_s$ が独立であることと $W_s \sim N(0, s)$, $W_t - W_s \sim W_{t-s} \sim N(0, t-s)$ を用いる。

(1) $f_{W_t}(x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}$. (2) $f_{(W_s, W_t - W_s)}(x, z) = \frac{1}{\sqrt{2\pi s}} e^{-\frac{x^2}{2s}} \frac{1}{\sqrt{2\pi(t-s)}} e^{-\frac{z^2}{2(t-s)}}$ より、

$$f_{(W_s, W_t)}(x, y) = \frac{1}{\sqrt{2\pi s}} e^{-\frac{x^2}{2s}} \frac{1}{\sqrt{2\pi(t-s)}} e^{-\frac{(y-x)^2}{2(t-s)}}.$$

(3) $E[W_s W_t] = E[W_s(W_t - W_s) + W_s^2] = E[W_s]E[W_t - W_s] + E[W_s^2] = s$.

(4) 独立増分性より、 $W_t - W_s$ と $W_u, 0 \leq u \leq s$ は独立。よって、

$$E[W_t - W_s + W_s | W_u, 0 \leq u \leq s] = E[W_t - W_s | W_u, 0 \leq u \leq s] + W_s = E[W_t - W_s] + W_s = W_s \quad \square$$

問題 3.7 W_t を標準ブラウン運動とする。次を求めよ。ただし、 $0 < s < t, \alpha \in \mathbf{R}$ とする。

- (a) $E[W_t^3]$, (b) $E[W_t^4]$, (c) $E[e^{\alpha W_t}]$, (d) $E[W_t e^{\alpha W_t}]$, (e) $E[W_s^2 W_t^2]$,
 (f) $E[W_t^2 | W_s]$, (g) $E[e^{\alpha W_t} | W_s]$, (h) $E[W_s e^{\alpha W_t}]$, (i) $P(W_t > 2W_s > 0)$

4 シミュレーション

確率変数 X の密度関数を $f(x)$ とすると、

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x) dx$$

であるが、この定積分を解析的に求めるのではなく、モンテカルロ法によって数値解析的に求める。

[モンテカルロ法] X_1, X_2, \dots を i.i.d. で $X_i \sim X$ とすると、 $g(X_1), g(X_2), \dots$ も i.i.d. で $g(X_i) \sim g(X)$ であり、大数の法則により

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n g(X_i) = E[g(X)]$$

となる。これより、 X の分布をもつ乱数 x_1, x_2, \dots を十分多く (N 個として) 発生できれば、

$$\int_{-\infty}^{\infty} g(x)f(x) dx \doteq \frac{1}{N} \sum_{i=1}^N g(x_i)$$

とできる。実際、 $U \sim U(0, 1)$ のとき $E[e^{U^2}] = \int_0^1 e^{x^2} dx$ を 10 個の $U(0, 1)$ に従う乱数

0.0623, 0.3246, 0.4265, 0.2796, 0.0424, 0.5218, 0.6877, 0.9943, 0.2771, 0.1668

を用いて (順に u_1, \dots, u_N , $N = 10$ と表す)、

$$E[e^{U^2}] = \int_0^1 e^{x^2} dx \doteq \frac{1}{N} \sum_{i=1}^N e^{u_i^2} \doteq 1.311$$

と求めることができる。別の 10 個の乱数で求めると 1.5859, 同様に 100 個の乱数で求めると 1.43676 となる (以上 Excel を用いた)。この積分の R で数値計算の値は 1.46265 である。

誤差評価や収束するスピードについて、中心極限定理より $\frac{\frac{1}{n} \sum_{i=1}^n g(X_i) - E[g(X)]}{\sqrt{\frac{V(g(X))}{n}}}$ は $N(0, 1)$ で近似できることに注意する。これより、 $\frac{1}{n} \sum_{i=1}^n g(X_i) - E[g(X)] = \sqrt{\frac{V(g(X))}{n}} N(0, 1) + o\left(\frac{1}{\sqrt{n}}\right)$ であり、もし $V(g(X))$ を小さくできれば、この誤差評価が良くなることがわかる (cf. 分散減少法)。

4.1 確率変数を生成する技法

4.1.1 逆関数法

一様分布に従う確率変数 U から $F(x)$ を分布関数とする確率変数 X を生成するには、

$$X = F^{\leftarrow}(U), \quad \text{ただし、} F^{\leftarrow}(u) = \inf\{x; F(x) \geq u\}$$

とすればよい。具体的には、連続型なら X のとり得る値となる x 上で $F(x)$ は連続で (狭義) 単調増加なので、 $F^{-1}(u) = F^{\leftarrow}(u)$ となり、 $X = F^{-1}(U)$ の分布関数は $F(x)$ となる。実際、

$$P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x).$$

また、 X が離散型で $P(X_k = x_k) = p_k$, $k = 1, 2, \dots, N$ とするとき、以下のように X を定めればよい。

$$\begin{aligned} 0 \leq U \leq p_1 & \longleftrightarrow X = x_1 \\ p_1 < U \leq p_1 + p_2 & \longleftrightarrow X = x_2 \\ p_1 + p_2 < U \leq p_1 + p_2 + p_3 & \longleftrightarrow X = x_3 \\ & \vdots \\ p_1 + p_2 + \dots + p_{N-1} < U \leq p_1 + p_2 + \dots + p_N = 1 & \longleftrightarrow X = x_N \end{aligned} \tag{4.1}$$

例 4.1 $X \sim \text{Ex}(1)$ に対し、これを一様分布 $U \sim U(0, 1)$ から逆関数法を用いて生成し、 $E[X]$ と $E[(1+X)^{-1}]$ をシミュレートする。

X の分布関数は $F(x) = 1 - e^{-x}$ ($x > 0$) であった。よって、 $F^{-1}(u) = -\log(1 - u)$ に注意して、 $X = -\log(1 - U)$ とおく。

次に左のように一様乱数 U が与えられたとする。 U から逆関数法で X を生成し、これから $(1 + X)^{-1}$ を計算したのが下の 2 行である。

U	0.84	0.30	0.76	0.35	0.16
X	1.8326	0.3566	1.4271	0.4307	0.1743
$(1 + X)^{-1}$	0.3530	0.7371	0.4120	0.6990	0.8516

例えば、 $U = 0.84$ に対しては、アクチュアリー資格試験 数学 付表 V 自然対数表を用いて計算すると、

$$X = -\log(1 - 0.84) = -\log \frac{1.6}{10} = 2.3026 - 0.4700 = 1.8326,$$

$$(1 + X)^{-1} = (1 + 1.8326)^{-1} = 0.35303 \dots$$

同様に、他の U に対しても上記の表のような $X, (1 + X)^{-1}$ を得る。

これより、 $E[X]$ をシミュレートすると、 $\frac{1}{5}(1.8326 + 0.3566 + 1.4271 + 0.4307 + 0.1743) = 0.84426$ 、

$E[(1 + X)^{-1}]$ については、 $\frac{1}{5}(0.3530 + 0.7371 + 0.4120 + 0.6990 + 0.8516) = 0.61054$ となる。□

問題 4.1 一様分布に従う確率変数 U から次の分布に従う確率変数 X を逆関数法により生成する。

右表の一様乱数 U から生成される X の実現値をそれぞれ求め、それにより平均値 $E[X]$ をシミュレートせよ。

	1	2	3	4	5
U	0.05	0.30	0.54	0.35	0.18
X	1	2	3	4	5
P	0.23	0.17	0.22	0.18	0.20

(1) 平均 2 の指数分布に従う確率変数 X 。

(2) 右の確率分布表で与えられる確率変数 X 。

問題 4.2 $0 < p < 1$ とする。 $U \sim U(0, 1)$ のとき、 $X = \text{int}\left(\frac{\log U}{\log(1 - p)}\right)$ とすると $X \sim \text{Ge}(p)$ を示せ。ここで、 $\text{int}(a)$ は a の整数部分を表す。次に、 $p = 0.2$ とし、問題 4.1 で与えられた一様乱数 U からこの式によって生成される X の実現値をそれぞれ求め、それにより平均値 $E[X]$ をシミュレートせよ。(注意 モデリングの教科書 pp.4-22-には、効率的な離散型確率変数に従う乱数の発生方法が述べられている。)

4.1.2 棄却法

密度関数 $g(y)$ である確率変数 Y から、密度関数 $f(x)$ である確率変数 X を生成したいとする。この際、

$$\text{ある定数 } c \text{ があって、} f(y)/g(y) \leq c \text{ for all } y \quad (4.2)$$

であるならば、次の手順により、 g に従う Y を用いて f に従う確率変数 X を求めることができる。

[手順 1] 密度関数が g である確率変数 Y を生成する。(例えば、逆関数法で生成する。)

[手順 2] $(0, 1)$ 上の一様分布に従う確率変数 U を生成する。

[手順 3] $U \leq \frac{f(Y)}{cg(Y)}$ ならば、 $X = Y$ と (X として Y を採用) し、 $U > \frac{f(Y)}{cg(Y)}$ ならば手順 1 にもどる。ここで c は (4.2) を満たす最小のものを選択することにより、計算手順数が効率化できることに注意する。

定理 4.1 棄却法で生成された X の密度関数は f 。

証明: $\forall y$ に対して $0 \leq \frac{f(y)}{cg(y)} \leq 1$ に注意する。

$$P(X \leq x) = P\left(Y \leq x \mid U \leq \frac{f(Y)}{cg(Y)}\right) = \frac{P\left(Y \leq x, U \leq \frac{f(Y)}{cg(Y)}\right)}{P\left(U \leq \frac{f(Y)}{cg(Y)}\right)} = \frac{\int_{-\infty}^x \int_0^{\frac{f(y)}{cg(y)}} g(y) \, du \, dy}{\int_{-\infty}^{\infty} \int_0^{\frac{f(y)}{cg(y)}} g(y) \, du \, dy}$$

$$= \frac{\int_{-\infty}^x \frac{f(y)}{cg(y)} g(y) dy}{\int_{-\infty}^{\infty} \frac{f(y)}{cg(y)} g(y) dy} = \frac{\frac{1}{c} \int_{-\infty}^x f(y) dy}{\frac{1}{c} \int_{-\infty}^{\infty} f(y) dy} = \int_{-\infty}^x f(y) dy. \quad \square$$

系 4.2 棄却法で、定理 4.1 の証明より $X = Y$ として採用される確率は $P\left(U \leq \frac{f(Y)}{cg(Y)}\right) = \frac{1}{c}$ である。ここで、 T で一つ採用されるまでの手順 1-3 の繰り返し回数を表すと、 $P(T = k) = \left(1 - \frac{1}{c}\right)^{k-1} \frac{1}{c}$, $k = 1, 2, \dots$, となり、平均繰り返し回数は

$$E[T] = \sum_{k=1}^{\infty} k \left(1 - \frac{1}{c}\right)^{k-1} \frac{1}{c} = \frac{\frac{1}{c}}{\left\{1 - \left(1 - \frac{1}{c}\right)\right\}^2} = c$$

となる。

注意 離散型の場合に棄却法で確率変数を生成する方法についてはモデリングの教科書 p.4-9 を参照のこと。

例 4.2 $Z \sim N(0, 1)$ に対し、 $X = |Z|$ を $Y \sim \text{Ex}(1)$ から棄却法を用いて生成する。

X の密度関数は $f(x) = \frac{2}{\sqrt{2\pi}} e^{-x^2/2}$ ($x > 0$) であった。よって、 Y の密度関数 $g(x) = e^{-x}$ ($x > 0$) で

$$\frac{f(x)}{g(x)} = \sqrt{\frac{2}{\pi}} e^{-x^2/2} = \sqrt{\frac{2}{\pi}} e^{-[(x-1)^2 - 1]/2} \leq \frac{f(1)}{g(1)} = \sqrt{\frac{2}{\pi}}.$$

よって、 $c = \sqrt{2e/\pi}$ ととり、 $\frac{f(x)}{cg(x)} = \exp\left(x - \frac{x^2}{2} - \frac{1}{2}\right) = \exp\left\{-\frac{(x-1)^2}{2}\right\}$. 従って、

[手順 1] 独立な $U_1, U_2 \sim U(0, 1)$ を準備する。

[手順 2] (Y の逆関数法による生成)

$F_Y(y) = 1 - e^{-y}$ より、 $F_Y^{-1}(u) = -\log(1 - u)$ に注意して、 $Y = -\log(1 - U_1)$ とおく。

[手順 3] $U_2 \leq \exp\left\{-(Y-1)^2/2\right\}$ なら $X = Y$ を採用。そうでなければ手順 1 にもどる。

次に左のように一様乱数 U_1, U_2 が与えられたとする。 U_1 から逆関数法で Y を生成し、さらに棄却法で X を生成する。

$F_Y(y) = 1 - e^{-y}$ より $F_Y^{-1}(u) = -\log(1 - u)$.

まず、 $Y = -\log(1 - U_1)$ を求め、記入する。

次に、 $Y' = e^{-(Y-1)^2/2}$ を求めた。例えば、 $U_1 = 0.84$ に対しては、

$$Y = -\log(1 - 0.84) = -\log \frac{1.6}{10} = 2.3026 - 0.4700 = 1.8326,$$

$$Y' = e^{-(1.8326-1)^2/2} = e^{-0.347\dots} = e^{-0.35} = (e^{-0.10})^3 e^{-0.05} = (0.9048)^3 \cdot 0.9512 = 0.7046$$

と、アクチュアリー資格試験 数学 付表 V 自然対数表および VI 指数関数表を用いて計算した*3。

同様に、他の U_1 に対しても上記の表のような Y, Y' を得る。最後に、 Y' と U_2 を比較して棄却されるデータは最後の一つとなり、 $E[X]$ をシミュレートすると、 $\frac{1}{4}(1.8326 + 0.3566 + 1.4271 + 0.4307) = 1.01175$ となる。 \square

問題 4.3 $X \sim \Gamma\left(\frac{3}{2}, 1\right)$ に対し、 $Y \sim \text{Ex}(\lambda)$ から棄却法を用いて生成する。

(1) 繰り返し回数が最小となる λ を求めよ。また、1 つ採用するための繰り返し回数の期待値を求めよ。

(2) (1) で求めた λ に対して、例 4.2 の一様乱数 U_1, U_2 から、逆関数法により Y を生成し、この Y から棄却法で X を生成する。これにより、 X の期待値をシミュレートせよ。

*3 $e^{0.01}, e^{-0.01}$ の値より Y' の値の誤差は 1% 程度となる。

問題 4.4 $f(x) = 30x(1-x)^4$, $0 < x < 1$ を密度関数にもつ確率変数を $Y \sim U(0, 1)$ から棄却法で生成する。

(1) 繰り返し回数を最小にする定数 c を定めよ。

(2) X を生成する手順を述べ、右表の Y および $U \sim U(0, 1)$ の

Y	0.05	0.30	0.54	0.35	0.18
U	0.08	0.77	0.69	0.57	0.77

シミュレーション結果から生成される X の平均値を求めよ。

4.1.3 合成法

F_1, \dots, F_n を分布関数で各々に対応する確率変数を生成する手段を持っているとする。このとき、 F がその重み付け平均

$$F(x) = \sum_{i=1}^n h_i F_i(x), \quad \sum_{i=1}^n h_i = 1, \quad h_i > 0$$

を分布関数とする確率変数 Y は次のようにシミュレーションする。(簡単のため $n = 3$ で述べる。)

独立な $U_1, U_2 \sim U(0, 1)$ を準備し、

$$\begin{aligned} 0 \leq U_1 \leq h_1 & \iff Y = F_1^{-1}(U_2) \\ h_1 < U_1 \leq h_1 + h_2 & \iff Y = F_2^{-1}(U_2) \\ h_1 + h_2 < U_1 \leq h_1 + h_2 + h_3 = 1 & \iff Y = F_3^{-1}(U_2) \end{aligned}$$

と Y を定める。ただし、 F_i が離散型の場合は (4.1) のように定義するものとする。

証明: $0 \leq h < k \leq 1$ に対し $P(h \leq U_1 \leq k, F_i^{-1}(U_2) \leq x) = P(h \leq U_1 \leq k)P(U_2 \leq F_i(x)) = (k-h)F_i(x)$ より

$$\begin{aligned} P(Y \leq x) &= P(0 \leq U_1 \leq h_1, F_1^{-1}(U_2) \leq x) + P(h_1 < U_1 \leq h_1 + h_2, F_2^{-1}(U_2) \leq x) \\ &\quad + P(h_1 + h_2 < U_1 \leq 1, F_3^{-1}(U_2) \leq x) \\ &= h_1 F_1(x) + h_2 F_2(x) + h_3 F_3(x). \quad \square \end{aligned}$$

例題 4.3 $f_X(x) = \begin{cases} 3x^3 + \frac{1}{2}x & (0 < x < 1) \\ 0 & (\text{その他}) \end{cases}$ となる確率変数 X を合成法を用いて作れ。

また、右の表の U_1, U_2 をシミュレーション結果として生成される X の平均値を求めよ。

U_1	0.23	0.94	0.78	0.55	0.20
U_2	0.94	0.56	0.62	0.92	0.91

解: $F_1(x) = x^4$, $F_2(x) = x^2$ とすると、 $F_X(x) = \frac{3}{4}F_1(x) + \frac{1}{4}F_2(x)$ ($0 < x < 1$) となることに注意する。

よって、 $0 \leq U_1 \leq 3/4$ なら $X = U_2^{1/4}$, $3/4 < U_1 \leq 1$ なら $X = U_2^{1/2}$ ととればよい。

$$U_1 = 0.23 \text{ について、} 0 \leq U_1 \leq 3/4 \text{ より、} X = F_1^{-1}(U_2) = (0.94)^{1/4} = 0.9846 \dots,$$

$$U_1 = 0.94 \text{ について、} 3/4 < U_1 \leq 1 \text{ より、} X = F_2^{-1}(U_2) = (0.56)^{1/2} = 0.7483 \dots,$$

同様にすると順に、 $X = (0.62)^{1/2} = 0.7874 \dots$, $X = (0.92)^{1/4} = 0.9793 \dots$, $X = (0.91)^{1/4} = 0.9766 \dots$ となり、 $E[X]$ をシミュレートすると、 $\frac{1}{5}(0.985 + 0.748 + 0.787 + 0.979 + 0.9766) = 0.853$ となる。 \square

問題 4.5 合成法に関する以下の問に答えよ。ただし、 X の実現値を求める際は、 $f_X(x)$ の右辺の左から順に、 U_1 の範囲を定めるものとし、 U_2 により逆関数法を用いて X の実現値を決定せよ。

(1) $f_X(x) = \begin{cases} \frac{1}{5}x^{-1/2} + \frac{4}{5}x + \frac{1}{5} & (0 < x < 1) \\ 0 & (\text{その他}) \end{cases}$ となる確率変数 X を合成法を用いて作れ。また、例題 4.3

の表の U_1, U_2 をシミュレーション結果として生成される X の平均値を小数点以下第 3 位まで求めよ。

(2) $g_1(x) = \begin{cases} 2(x+1)^{-3} & (x > 0) \\ 0 & (\text{その他}) \end{cases}$ とし、 $g_2(x)$ を $[0, 5]$ 上の一様分布の密度関数とする。

$f_X(x) = \frac{2}{3}g_1(x) + \frac{1}{3}g_2(x)$ となる確率変数 X を合成法を用いて作れ。また、例題 4.3 の表の U_1, U_2 をシミュレーション結果として生成される X の平均値を小数点以下第 2 位まで求めよ。

(3) $g_1(x) = \begin{cases} 0.5e^{-0.5x} & (x > 0) \\ 0 & (\text{その他}) \end{cases}$, $g_2(x) = \begin{cases} e^x & (x < 0) \\ 0 & (\text{その他}) \end{cases}$ とする。 $f_X(x) = \frac{2}{3}g_1(x) + \frac{1}{3}g_2(x)$ となる

確率変数 X を合成法を用いて作れ。また、 $(U_1, U_2) = (0.32, 0.75), (0.81, 0.35)$ として X の実現値をそれぞれ求めよ。

4.2 分散減少法

この章の最初に述べたが、 X_1, X_2, \dots を i.i.d. とし、 $g(X_i)$ の平均を θ 、標準偏差を σ とする。このとき、その標本平均 $\hat{\theta}_n = \frac{1}{n} \sum_{i=1}^n g(X_i)$ は、 $E[\hat{\theta}_n] = E[g(X_1)] = \theta$ であり、

$$V(\hat{\theta}_n) = \frac{1}{n^2} \sum_{i=1}^n V(g(X_i)) = \frac{\sigma^2}{n}, \quad \text{すなわち} \quad E[(\hat{\theta}_n - \theta)^2] = \frac{\sigma^2}{n}.$$

これより、 $\hat{\theta}_n$ と θ の標準誤差 (二乗平均誤差の平方根) は c/\sqrt{n} という形となる。これは n をシナリオ数とすると精度 c/\sqrt{n} となる。即ち、1000 本のシナリオを用いたモンテカルロ法により得られた値の精度をもう一ケタ上げるためには、その 100 倍である 10 万本のシナリオが必要となる。

ここではこの c を減少させることにより精度をよくする方法である分散減少法を紹介する。(詳しくはモデリングの教科書 pp.4-31-を参照のこと。)

4.2.1 負の相関法

定理 4.3 f, g がともに単調増加もしくはともに単調減少なら、 $\text{Cov}(f(X), g(X)) \geq 0$.

証明: $x \leq y$ なら $(f(x) - f(y))(g(x) - g(y)) \geq 0$, $x \geq y$ でも $(f(x) - f(y))(g(x) - g(y)) \geq 0$ より、常に $(f(X) - f(Y))(g(X) - g(Y)) \geq 0$ は成立する。ここで、 X, Y が独立で同じ分布に従うとすると、

$$\begin{aligned} 0 &\leq E[(f(X) - f(Y))(g(X) - g(Y))] = E[f(X)g(X)] - E[f(X)g(Y)] - E[f(Y)g(X)] + E[f(Y)g(Y)] \\ &= E[f(X)g(X)] - E[f(X)]E[g(Y)] - E[f(Y)]E[g(X)] + E[f(Y)g(Y)] \quad (\because X, Y \text{ は独立}) \\ &= E[f(X)g(X)] - E[f(X)]E[g(X)] - E[f(X)]E[g(X)] + E[f(X)g(X)] \quad (\because X, Y \text{ は同分布}) \\ &= 2(E[f(X)g(X)] - E[f(X)]E[g(X)]) = 2\text{Cov}(f(X), g(X)). \quad \square \end{aligned}$$

これより、 $U \sim U(0, 1)$ とし、 $h(x)$ が単調増加 (減少) なら、 $-h(1-x)$ も単調増加 (減少) なので、

$$\text{Cov}(h(U), h(1-U)) = -\text{Cov}(h(U), -h(1-U)) \leq 0$$

となる。

X の分布関数を F とし、その逆関数 F^{-1} が存在するとする。 $U_i, i = 1, 2, \dots$ を $U(0, 1)$ に従う i.i.d. とすると、 $1 - U_i \sim U(0, 1)$ なので、

$$X_i = F^{-1}(U_i), \quad Y_i = F^{-1}(1 - U_i)$$

とおくと、分布 F に従う確率変数 X_i, Y_i を生成できる。ここで、

$$\hat{\theta}_{2M}^{(1)} = \frac{1}{2M} \sum_{i=1}^M \{g(X_i) + g(Y_i)\} = \frac{1}{2M} \sum_{i=1}^M \{g(F^{-1}(U_i)) + g(F^{-1}(1 - U_i))\}$$

とすると、これは $E[\hat{\theta}_{2M}^{(1)}] = \theta$ を満たす。ただし、 θ は $g(X)$ の平均であった。ここで、もし $g(x)$ が単調なら、 $g(F^{-1}(u))$ も単調なので、 $\text{Cov}(g(F^{-1}(U_i)), g(F^{-1}(1 - U_i))) \leq 0$ となることに注意すると、

$$\begin{aligned} V(\hat{\theta}_{2M}^{(1)}) &= \frac{1}{(2M)^2} \sum_{i=1}^M V[g(F^{-1}(U_i)) + g(F^{-1}(1 - U_i))] \\ &= \frac{1}{(2M)^2} \sum_{i=1}^M \left\{ V[g(F^{-1}(U_i))] + 2\text{Cov}(g(F^{-1}(U_i)), g(F^{-1}(1 - U_i))) + V[g(F^{-1}(1 - U_i))] \right\} \end{aligned}$$

$$\leq \frac{1}{(2M)^2} \sum_{i=1}^M \{V[g(F^{-1}(U_i))] + V[g(F^{-1}(1-U_i))]\} = \frac{1}{2M} V(g(X)) = V(\hat{\theta}_{2M}).$$

ここで、3行目の最初の等号は $F^{-1}(U_i)$, $F^{-1}(1-U_i)$ とともに X と同じ分布に従うことを用いた。以上より、 $\hat{\theta}_{2M}^{(1)}$ は θ の不偏推定量であり、標本平均 $\hat{\theta}_{2M}$ より分散が小さくなることがわかった。また、一様分布に従う確率変数を $2M$ 個生成する必要はなく、 M 個のみ生成すればよいので、一様分布に従う確率変数を生成する時間も節約できる。

例題 4.4 $\int_0^1 \sin \frac{\pi}{2} x dx$ を、(a) $2M$ 個の一様乱数 U_1, U_2, \dots, U_{2M} を用いてモンテカルロシミュレーションをした場合の誤差の分散 $V(\hat{\theta}_{2M})$ を求めよ。

(b) $2M$ 個の一様乱数 $U_1, U_2, \dots, U_M, 1-U_1, 1-U_2, \dots, 1-U_M$ を用いた誤差の分散 $V(\hat{\theta}_{2M}^{(1)})$ を求めよ。

(c) 問題 4.1 の一様乱数 U を用いて、 $\hat{\theta}_M$ と $\hat{\theta}_{2M}^{(1)}$, $M = 5$, の実現値を求めよ。

解: 以下、独立性を考えない式では U_i を U と表す。 $g(x) = \sin \frac{\pi}{2} x$ とおく。 $\hat{\theta}_{2M} = \frac{1}{2M} \sum_{i=1}^{2M} g(U_i)$ について、

$$E[g(U_i)] = \int_0^1 \sin \frac{\pi}{2} x dx = \frac{2}{\pi}, \quad E[g(U_i)^2] = \int_0^1 \sin^2 \frac{\pi}{2} x dx = \frac{1}{2} \quad \text{より}$$

$$V(\hat{\theta}_{2M}) = \frac{1}{(2M)^2} \sum_{i=1}^{2M} V(g(U_i)) = \frac{1}{2M} V(g(U)) = \frac{1}{2M} \left(\frac{1}{2} - \frac{4}{\pi^2} \right) \doteq \frac{1}{2M} 0.09471565 \dots$$

$\hat{\theta}_{2M}^{(1)} = \frac{1}{2M} \sum_{i=1}^M \{g(U_i) + g(1-U_i)\}$ について、

$$\begin{aligned} E[\{g(U_i) + g(1-U_i)\}^2] &= E[g(U_i)^2] + E[g(1-U_i)^2] + 2E[g(U_i)g(1-U_i)] \\ &= \frac{1}{2} + \frac{1}{2} + \int_0^1 2 \sin \frac{\pi}{2} x \sin \frac{\pi}{2} (1-x) dx = 1 + \int_0^1 2 \sin \frac{\pi}{2} x \cos \frac{\pi}{2} x dx = 1 + \frac{2}{\pi} \quad \text{より} \end{aligned}$$

$$\begin{aligned} V(\hat{\theta}_{2M}^{(1)}) &= \frac{1}{(2M)^2} \sum_{i=1}^M V(g(U_i) + g(1-U_i)) = \frac{1}{4M} V(g(U) + g(1-U)) \\ &= \frac{1}{4M} [E[\{g(U) + g(1-U)\}^2] - (E[g(U) + g(1-U)])^2] \\ &= \frac{1}{4M} \left(1 + \frac{2}{\pi} - \frac{16}{\pi^2} \right) \doteq \frac{1}{4M} 0.007740417054 \dots \end{aligned}$$

右表のより、 $\hat{\theta}_5 = \frac{2.084}{5} = 0.417$,

$$\hat{\theta}_{2.5}^{(1)} = \frac{2.084 + 4.362}{10} = 0.645. \quad \square$$

U	0.05	0.30	0.54	0.35	0.18	計
$\sin \frac{\pi}{2} U$	0.078	0.454	0.750	0.522	0.279	2.084
$\sin \frac{\pi}{2} (1-U)$	0.997	0.891	0.661	0.853	0.960	4.362

問題 4.6 $\int_0^1 \sqrt{x} dx$ を、(a) $2M$ 個の一様乱数 U_1, U_2, \dots, U_{2M} を用いてモンテカルロシミュレーションをした場合の誤差の分散 $V(\hat{\theta}_{2M})$ を求めよ。

(b) $2M$ 個の一様乱数 $U_1, U_2, \dots, U_M, 1-U_1, 1-U_2, \dots, 1-U_M$ を用いた誤差の分散 $V(\hat{\theta}_{2M}^{(1)})$ を求めよ。

(c) 問題 4.1 の一様乱数 U を用いて、 $\hat{\theta}_M$ と $\hat{\theta}_{2M}^{(1)}$, $M = 5$, の実現値を求めよ。

4.2.2 制御変量法

$\hat{\theta}_n^{(3)} = \frac{1}{n} \sum_{i=1}^n \{g(X_i) + c(h(X) - \theta_h)\}$ とし、 $E[h(X_i)] = \theta_h$ は既知と仮定すると (この h を制御変量という)、

$$V(\hat{\theta}_n^{(3)}) = \frac{1}{n} V[g(X) + c(h(X) - \theta_h)] = \frac{1}{n} \{V(g(X)) + 2c \text{Cov}(g(X), h(X)) + c^2 V(h(X))\}$$

$$= \frac{1}{n} \left[V(h(X)) \left\{ c + \frac{\text{Cov}(g(X), h(X))}{V(h(X))} \right\}^2 + V(g(X)) - \frac{(\text{Cov}(g(X), h(X)))^2}{V(h(X))} \right] \quad (4.3)$$

より、 $c = c^* = -\frac{\text{Cov}(g(X), h(X))}{V(h(X))}$ ととると、分散は最小となり、このとき

$$V(\hat{\theta}_n^{(3)}) = \frac{1}{n} \left\{ V(g(X)) - \frac{(\text{Cov}(g(X), h(X)))^2}{V(h(X))} \right\} \leq \frac{1}{n} V(g(X)) = V(\hat{\theta}_n) \quad \text{となる。}$$

ここで、 $\frac{V(\hat{\theta}_n^{(3)})}{V(\hat{\theta}_n)} = 1 - \frac{(\text{Cov}(g(X), h(X)))^2}{V(h(X))V(g(X))} = 1 - (\text{Corr}(g(X), h(X)))^2$,

つまり分散の減少率は $g(X)$ と $h(X)$ の相関係数 $\text{Corr}(g(X), h(X))$ の 2 乗となる。

一方、標本値だけでは、

$$\widehat{\text{Cov}}[g(X), h(X)] = \frac{1}{n-1} \sum_{i=1}^n (g(X_i) - \hat{g}(X))(h(X_i) - \hat{h}(X)),$$

$$\widehat{\text{Var}}[h(X)] = \frac{1}{n-1} \sum_{i=1}^n (h(X_i) - \hat{h}(X))^2, \quad \text{ここで、} \hat{g}(X) = \frac{1}{n} \sum_{i=1}^n g(X_i), \quad \hat{h}(X) = \frac{1}{n} \sum_{i=1}^n h(X_i).$$

c を $c^* = -\frac{\widehat{\text{Cov}}[g(X), h(X)]}{\widehat{\text{Var}}[h(X)]}$ で推定する。

以上より、 $\hat{\theta}_n^{(3)} = \hat{g}(X) - \frac{\widehat{\text{Cov}}[g(X), h(X)]}{\widehat{\text{Var}}[h(X)]} (\hat{h}(X) - \theta_h)$ がその実現値となる。

例題 4.5 $\int_0^1 e^x dx$ を、(a) n 個の一樣乱数 U_1, U_2, \dots, U_n を用いてモンテカルロシミュレーションをした場合の誤差の分散 $V(\hat{\theta}_n)$ を求めよ。(b) 制御変量 $h(U) = U$ をとったときの誤差の分散 $V(\hat{\theta}_n^{(3)})$ を計算せよ。(c) 問題 4.1 の一樣乱数 U の標本値を用いて、 $\hat{\theta}_n$ と $\hat{\theta}_n^{(3)}$, $n = 5$, の実現値を求めよ。

解: $g(U) = e^U$ とおく。 $\hat{\theta}_n = \frac{1}{n} \sum_{i=1}^n g(U_i)$ について、

$$E[g(U_i)] = \int_0^1 e^u du = e - 1, \quad E[g(U_i)^2] = \int_0^1 e^{2u} du = \frac{e^2 - 1}{2} \quad \text{より}$$

$$V(\hat{\theta}_n) = \frac{1}{n} V(g(U)) = \frac{1}{n} \left(\frac{e^2 - 1}{2} - (e - 1)^2 \right) = \frac{1}{n} \frac{(e - 1)(3 - e)}{2} \doteq \frac{1}{n} 0.242035607 \dots$$

$\hat{\theta}_n^{(3)} = \frac{1}{n} \sum_{i=1}^n \{g(U_i) + c(h(U_i) - \theta_h)\}$ について、(4.3) より $c = -\frac{\text{Cov}(g(U), h(U))}{V(h(U))}$ とすれば、

$$V(h(U)) = E[U^2] - (E[U])^2 = \frac{1}{12}, \quad \text{Cov}(g(U), h(U)) = \int_0^1 x e^x dx - \frac{1}{2} \int_0^1 e^x dx = 1 - \frac{e - 1}{2} \quad \text{より}$$

$$\begin{aligned} V(\hat{\theta}_n^{(3)}) &= \frac{1}{n} \left[V(g(U)) - \frac{(\text{Cov}(g(U), h(U)))^2}{V(h(U))} \right] = \frac{1}{n} \left[\frac{e^2 - 1}{2} - (e - 1)^2 - 12 \left(1 - \frac{e - 1}{2} \right)^2 \right] \\ &= \frac{1}{n} \frac{(3 - e)(7e - 19)}{2} = \frac{1}{n} 0.0394022292 \dots \end{aligned}$$

標本値では、 $\widehat{\text{Var}}(U) = 0.03393$, $\widehat{\text{Cov}}[e^U, U] = 0.046032$ より、 $\hat{c}^* = -1.3567$, $\theta_h = 0.5$ (既知) より次の表を得る。よって、 $\hat{\theta}_5 = \frac{6.733}{5} = 1.3466$,

$$\hat{\theta}_5^{(3)} = \hat{\theta}_5 + \hat{c}^* (\hat{h}(U) - \theta_h)$$

$$= 1.3466 - 1.3567(0.284 - 0.5) = 1.6396. \quad \square$$

U	0.05	0.30	0.54	0.35	0.18	計
e^U	1.051	1.350	1.716	1.419	1.197	6.733

問題 4.7 $\int_0^1 \sqrt{1 - x^2} dx$ を (a) n 個の一樣乱数 U_1, U_2, \dots, U_n を用いてモンテカルロシミュレーションをした場合の誤差の分散 $V(\hat{\theta}_n)$ を求めよ。

- (b) 制御変量 $h(U) = U$ をとったときの誤差の分散 $V(\hat{\theta}_n^{(3)})$ を計算せよ。
(c) 制御変量 $h(U) = U^2$ をとったときの誤差の分散 $V(\hat{\theta}_n^{(3)})$ を計算せよ。
(d) 問題 4.1 の一様乱数 U を用いて、 $\hat{\theta}_5$ と (b) の $\hat{\theta}_5^{(3)}$ の実現値を求めよ。

4.3 統計の復習 3 適合度、独立性の検定

4.3.1 適合度の検定

n 個のデータが d 個の階級 A_1, \dots, A_d に分類できたとき、 A_j が出現する母比率を p_j とし、次を考える。

$$\text{帰無仮説 } H_0: (p_1, \dots, p_d) = (p_1^0, \dots, p_d^0), \quad \text{対立仮説 } H_1: (p_1, \dots, p_d) \neq (p_1^0, \dots, p_d^0).$$

このとき、 A_j の出現回数を n_j とし、統計量

$$T_n = \sum_{j=1}^d \frac{(n_j - np_j^0)^2}{np_j^0} = \sum \frac{(\text{観察度数} - \text{期待度数})^2}{\text{期待度数}} \quad (4.4)$$

を考える。このとき、 H_0 のもと T_n は自由度 $d-1$ の χ^2 分布に法則収束する (確率統計学 II で示す) から、もし有意水準 α で検定するのであれば、(4.4) に実現値を代入した値を t とするとき、

$$\begin{aligned} \text{もし } t < \chi_{d-1}^2(\alpha) \text{ であれば } H_0 \text{ を採択し、} \\ \text{もし } t \geq \chi_{d-1}^2(\alpha) \text{ であれば } H_0 \text{ を棄却する。} \end{aligned}$$

ただし、 $np_j^0 < 5$ となる階級があるときは、隣接の階級と合併しすべての階級で $np_j^0 \geq 5$ となるように分けな
おす。

例題 4.6 ある県の成人を母集団とし、無作為抽出された 1000 人の血液型を調べたところ下の表のような観測値を得た。この県の血液型の分布は日本人の血液型の分布と同じであるといえるか。有意水準 5% で検定せよ。ただし、日本人の血液型の分布は $A : B : O : AB = 38.0 : 21.8 : 30.8 : 9.4$ である。

血液型	A	B	O	AB
観測値	360	216	330	94

解: A, B, O, AB の比率を順に p_1, p_2, p_3, p_4 とし、次の仮説を設定する。

$$H_0 : (p_1, p_2, p_3, p_4) = (0.380, 0.218, 0.308, 0.094), \quad H_1 \text{ は } H_0 \text{ の否定.}$$

有意水準 5% であるから、棄却域は自由度が $4-1$ であることに注意して

$$T \geq \chi_3^2(0.05) = 7.8147$$

である。期待度数は A, B, O, AB の順に 380, 218, 308, 94 であらから、実現値を代入して、

$$t = \frac{(360 - 380)^2}{380} + \frac{(216 - 218)^2}{218} + \frac{(330 - 308)^2}{308} + \frac{(94 - 94)^2}{94} = 2.6424 \dots$$

この値は棄却域に入らないから、 H_0 は採択される。従って、この県の血液型の分布は日本人の血液型の分布と同じといえる。 \square

問題 4.8 ある国の成人を母集団とし、無作為抽出された 100 人の血液型を調べたところ下の表のような観測値を得た。この県の血液型の分布は日本人の血液型の分布と同じであるといえるか。有意水準 5% で検定せよ。ただし、日本人の血液型の分布は $A : B : O : AB = 38.0 : 21.8 : 30.8 : 9.4$ である。

血液型	A	B	O	AB
観測値	28	23	42	7

問題 4.9 次のデータは 4 枚のコインを同時に投げ表の枚数を 100 回行い記録したものである。

これは二項分布 $B(4, 1/2)$ に従っているといえるか。有意水準 5% で検定せよ。

表の枚数	0	1	2	3	4
観測値	5	23	33	30	9

4.3.2 独立性の検定

N 個のデータが 2 種類の属性 A, B によるそれぞれの各階級 A_1, A_2, \dots, A_r および B_1, B_2, \dots, B_s に分割されて、右の表のような度数水河できたものとする。これを $r \times s$ 分割表という。ここで、 $A_i \cap B_j$ なる性質をもったデータ数は f_{ij} とする。このとき、次を考える。

帰無仮説 H_0 : 二つの属性 A, B は独立である

ことを有意水準 α で検定するには、すべての度数 $f_{ij} \geq 5$ かつ $f_{i\bullet} \cdot f_{\bullet j} / N \geq 5$ であるとき、

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^s \frac{\left(f_{ij} - \frac{f_{i\bullet} \cdot f_{\bullet j}}{N}\right)^2}{\frac{f_{i\bullet} \cdot f_{\bullet j}}{N}} > \chi_{\phi}^2(\alpha) \text{ ならば } H_0 \text{ を棄却し、}$$

$\chi^2 < \chi_{\phi}^2(\alpha)$ ならば H_0 を採択することにすればよい。ここで $\phi = (r-1)(s-1)$ である。

	B				
	B_1	B_2	\cdots	B_s	計
A					
A_1	f_{11}	f_{12}	\cdots	f_{1s}	$f_{1\bullet}$
A_2	f_{21}	f_{22}	\cdots	f_{2s}	$f_{2\bullet}$
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
A_r	f_{r1}	f_{r2}	\cdots	f_{rs}	$f_{r\bullet}$
計	$f_{\bullet 1}$	$f_{\bullet 2}$	\cdots	$f_{\bullet s}$	N

注意 4.1 (1) H_0 が正しいとして分割表を作れば、 $A_i \cap B_j$ のクラスに入るデータ数の期待値は $f_{i\bullet} \cdot f_{\bullet j} / N$ である。従って、統計量 χ^2 は

$$\chi^2 = \sum_i \sum_j \frac{(\text{実現値} - \text{期待度数})^2}{\text{期待度数}}$$

と (4.4) と同様となる。ただし、自由度は $\phi = (r-1)(s-1)$ である。

(2) 2×2 分割表の簡便計算法: 2×2 分割表のときには、上の χ^2 の値は次のようにして計算すると便利である (第 2 の等式が成立することを確かめよ)。

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^s \frac{\left(f_{ij} - \frac{f_{i\bullet} \cdot f_{\bullet j}}{N}\right)^2}{\frac{f_{i\bullet} \cdot f_{\bullet j}}{N}} = \frac{(f_{11}f_{22} - f_{12}f_{21})^2 N}{f_{1\bullet}f_{2\bullet}f_{\bullet 1}f_{\bullet 2}}. \quad (4.5)$$

(3) 2×2 分割表における **Yates の修正法**: もし、(2) の χ^2 の計算のときに、 f_{ij} の中に 5 より小さな値をとるものがあつたなら

$$\chi^2 = \frac{\left(|f_{11}f_{22} - f_{12}f_{21}| - \frac{N}{2}\right)^2 N}{f_{1\bullet}f_{2\bullet}f_{\bullet 1}f_{\bullet 2}},$$

とすればよい。このときも自由度はもちろん $(2-1) \times (2-1) = 1$ である。

(4) **Fisher の直接確率計算法**: (3) の検定において、度数 f_{ij} が小さすぎるとき χ^2 分布は利用できない。 H_0 : A, B は独立である という仮説の下で、全ての周辺度数 $f_{i\bullet}, f_{\bullet j}$ が一定のときの $r \times s$ 分割表のように各 $A_i \cap B_j$ の度数が f_{ij} となる確率が $\frac{\prod_{i=1}^r f_{i\bullet}! \prod_{j=1}^s f_{\bullet j}!}{N! \prod_{i=1}^r \prod_{j=1}^s f_{ij}!}$ であることを用い、実現値以上に偏った度数分布になるすべての場合の確率の和 p を計算し、有意水準と比較する。

例題 4.7 ある都市で有権者の A 内閣に対する支持率を調べた。

有権者から男性 150 人、女性 100 人を抽出し、支持する、しないを調べたらその人数は右の表のようであった。A 内閣の支持率は男性と女性とで、違いがあるとみてよいか、有意水準 5% で検定せよ。

	男	女	計
支持する	75	60	135
支持しない	75	40	115
計	150	100	250

解: 次の仮説を設定する。 H_0 : 男性女性と支持するしないは関係ない、 H_1 は H_0 の否定。

有意水準 5% であるから、棄却域は自由度が $(2-1)(2-1) = 1$ であることに注意して

$$T \geq \chi_1^2(0.05) = 3.8415$$

である。2 × 2 分割表の簡便計算法を用いて実現値を代入すると、

$$\chi^2 = \frac{(75 \cdot 40 - 60 \cdot 75)^2 \cdot 250}{135 \cdot 115 \cdot 150 \cdot 100} = 2.415 \dots$$

これは棄却域に入らないから、 H_0 は採択される。従って、男性女性と支持不支持には関係がないといえる。□

問題 4.10 (1) ある病気の予防注射の効果を調べるために、300 人を調査したところ、下の表 (1) の結果を得た。予防注射の効果があるといえるか。有意水準 5% で検定せよ。

(2) 下の表 (2) は 300 人の自動車所有者を年齢と過去 2 年間に起こした事故数に応じて分類したものである。年齢と事故数の間に関係があるかどうかを有意水準 5% で検定せよ。

表 (1)

	発病	発病なし
予防注射あり	38	142
予防注射なし	58	62

表 (2)

事故数	0	1 ~ 2	3
21 歳以下	8	23	14
22 ~ 26	21	42	12
27 歳以上	71	90	19

5 損保数理に関する確率統計の話題から

5.1 最尤推定量の漸近挙動

ここでは、前期に学んだ最尤推定量に関する、極限定理を扱う。次を参照した。

[LC] E.L. Lehmann, G. Casella: Theory of Point Estimation, Second Edition, Springer, 1998

確率統計学 II で学ぶ次の定義と定理を述べておく。この節ではこれが基本となる。

定義 5.1 (1) $\forall \delta > 0$ に対して $\lim_{n \rightarrow \infty} P(|X_n - X| < \delta) = 1$ となるとき、 X_n は X に確率収束するといひ、 $X_n \rightarrow X$ in prob. と表す。

(2) 確率変数 Y の分布関数を $F_Y(y) = P(Y \leq y)$ と表すとする。 X_n が X に法則収束するとは、 F_X の任意の連続点 c に対して、 $\lim_{n \rightarrow \infty} F_{X_n}(c) = F_X(c)$ となるときにいう。このとき、 $X_n \rightarrow X$ in law と表す。

定理 5.1 X_1, X_2, \dots を i.i.d. とし、 $S_n = \sum_{i=1}^n X_i$ とする。

(大数の法則) $\lim_{x \rightarrow \infty} xP(|X_1| > x) = 0$ ならば、数列 a_n があって $\frac{1}{n}S_n - a_n \rightarrow 0$ in prob. となる。特に、 $E[|X_1|] < \infty$ であれば、 $\frac{1}{n}S_n \rightarrow E[X_1]$ in prob. となる。(正確には大数の弱法則という。)

(中心極限定理) $E[X_1^2] < \infty$ とする。 $\mu = E[X_1]$, $\sigma^2 = V(X_1)$ とすると、 $\frac{S_n - n\mu}{\sqrt{n}} \rightarrow N(0, \sigma^2)$ in law とする。

θ は未知母数とし Θ をその母数空間とする。 X_1, X_2, \dots, X_n は i.i.d. でその密度関数を $f(x|\theta)$ あるいは確率関数を $p(x|\theta) = P_\theta(X = x)$ とする。以下、連続型の場合を考え、その確率を P_θ , 期待値を E_θ で表すものとする。(離散型の場合も同様に示せる。) さらに次を仮定する。

(A1) $\theta \neq \theta'$ なら密度関数として $f(x|\theta) \neq f(x|\theta')$ 。

(A2) $A = \{x | f(x|\theta) > 0\}$ は θ によらない。

(A3) 母数空間 Θ は開集合とし、真のパラメータ θ_0 はその内点であるとする。

このとき、尤度関数を $L_n(\theta|\mathbf{x}) = \prod_{i=1}^n f(x_i|\theta)$, 対数尤度を $l_n(\theta|\mathbf{x}) = \sum_{i=1}^n \log f(x_i|\theta)$ と表す。ただし、 $\mathbf{x} = (x_1, x_2, \dots)$ とする。この尤度関数もしくは対数尤度を最大にする $\hat{\theta}_n = \hat{\theta}_n(x_1, \dots, x_n)$:

$$L_n(\hat{\theta}_n|\mathbf{x}) = \max_{\theta \in \Theta} L_n(\theta|\mathbf{x}) \quad \text{または} \quad l_n(\hat{\theta}_n|\mathbf{x}) = \max_{\theta \in \Theta} l_n(\theta|\mathbf{x})$$

に対して、 $\hat{\theta}_n = \hat{\theta}_n(X_1, \dots, X_n)$ が最尤推定量であった。

定理 5.2 (A1)–(A3) の条件のもと、 $\forall \theta \neq \theta_0$ に対して、

$$P_{\theta_0}(L_n(\theta_0|\mathbf{X}) > L_n(\theta|\mathbf{X})) \rightarrow 1, \quad \text{as } n \rightarrow \infty,$$

ただし $\mathbf{X} = (X_1, X_2, \dots)$ 。

証明: $L_n(\theta_0|\mathbf{X}) > L_n(\theta|\mathbf{X}) \iff \frac{1}{n} \sum_{i=1}^n \log \frac{f(X_i|\theta)}{f(X_i|\theta_0)} < 0$ に注意する。ここで、大数の法則により

$$\frac{1}{n} \sum_{i=1}^n \log \frac{f(X_i|\theta)}{f(X_i|\theta_0)} \rightarrow E_{\theta_0} \left[\log \frac{f(X|\theta)}{f(X|\theta_0)} \right] \quad \text{in prob.}$$

とできる。ここで、Jensen の不等式を $(\log x)'' < 0$ に注意して用いると、

$$E_{\theta_0} \left[\log \frac{f(X|\theta)}{f(X|\theta_0)} \right] < \log E_{\theta_0} \left[\frac{f(X|\theta)}{f(X|\theta_0)} \right] = \log \int_{-\infty}^{\infty} \frac{f(x|\theta)}{f(x|\theta_0)} f(x|\theta_0) dx = 0$$

となり主張を得る。 □

命題 5.3 (Jensen の不等式) $f(x)$ が下に凸であれば

$$E[f(X)] \geq f(E[X]).$$

特に、 $f''(x) > 0$ であれば等号成立は X が定数のときに限る。

証明: $\mu = E[X]$ とする。 $f(x)$ は下に凸であるから $c \in \mathbf{R}$ があって $f(x) \geq c(x - \mu) + f(\mu)$ とできる。従って、

$$E[f(X)] \geq E[c(X - \mu) + f(\mu)] = c(E[X] - \mu) + f(\mu) = f(\mu).$$

$f''(x) > 0$ であれば $f(x) > f'(\mu)(x - \mu) + \mu$, $x \neq \mu$, となるので、 $P(X = \mu) = 1$ でなければ等号は成立しない。 □

定理 5.4 (最尤推定量の一致性) (A1)–(A3) の条件のもと、 $\forall x$ に対し $\Theta \ni \theta \mapsto f(x|\theta)$ は微分可能で θ についての偏微分を $f'(x|\theta)$ と表すとき、尤度方程式

$$l'(\theta|\mathbf{x}) = \sum_{i=1}^n \frac{f'(x_i|\theta)}{f(x_i|\theta)} = 0 \quad (5.1)$$

は根 $\hat{\theta}_n = \hat{\theta}_n(x_1, \dots, x_n)$ で $\hat{\theta}_n(X_1, \dots, X_n) \rightarrow \theta_0$ in prob. となるものをもつ。特に、最尤方程式 (5.1) がただ一つの解 $\hat{\theta}_n$ をもつ、すなわち、 $\hat{\theta}_n$ が最尤推定量であれば、 $\hat{\theta}_n$ は真のパラメータ θ_0 に確率収束する。すなわち、最尤推定量は一致推定量である。

証明: $a > 0$ を $[\theta_0 - a, \theta_0 + a] \subset \Theta$ となるようにとり、

$$S_n = \{ \mathbf{x}; l_n(\theta_0|\mathbf{x}) > l_n(\theta_0 - a|\mathbf{x}) \text{ and } l_n(\theta_0|\mathbf{x}) > l_n(\theta_0 + a|\mathbf{x}) \}$$

とする。定理 5.2 より $P_{\theta_0}(\mathbf{X} \in S_n) \rightarrow 1$ 。ここで、 $\forall \mathbf{x} \in S_n$ に対して、 $\hat{\theta}_n \in (\theta_0 - a, \theta_0 + a)$ をそこで $l_n(\theta|\mathbf{x})$ が極大となるようにとれる。ここで $l'(\hat{\theta}_n|\mathbf{x}) = 0$ に注意する。よって、

$$P_{\theta_0}(|\hat{\theta}_n - \theta_0| < a) \geq P_{\theta_0}(\mathbf{X} \in S_n) \rightarrow 1.$$

ここで、上記の $\hat{\theta}_n$ が a に依存していることに注意する。これを解決するためには、 $l_n(\theta|\mathbf{x})$ が極大となる θ で $|\theta - \theta_0|$ を最小とするものを選び $\hat{\theta}_n$ とすればよい。 $(\mathbf{x} \mapsto \hat{\theta}_n)$ の可測性が問題となるが、その証明は省略する。(Lehmann-Casella [LC] p.504 を参照のこと。) □

漸近有効性のために次の命題を準備する。

命題 5.5 確率変数 X_n が定数 $a > 0$ に確率収束し、確率変数 Z_n が Z に法則収束するとき、 Z_n/X_n は Z/a に法則収束する。

証明: c が $F_Z(z)$ が連続点 $\iff c/a$ が $F_{Z/a}(z/a)$ が連続点であり、 $F_Z(z) = F_{Z/a}(z/a)$ より、 F_Z の任意の連続点 c に対して、 $\lim_{n \rightarrow \infty} P(Z_n/X_n \leq c/a) = P(Z \leq c)$ を示せばよい。 $\varepsilon > 0$ を任意にとっておく。 $\delta > 0$ を

$$|x - c| < \delta \implies |F_Z(x) - F_Z(c)| < \frac{\varepsilon}{4}$$

ととる。 F_Z の不連続点は高々可算個なので、 c_1, c_2 を F_Z の連続点で $c - \delta < c_1 < c < c_2 < c + \delta$ となるようにできる。次に $\lim_{n \rightarrow \infty} F_{Z_n}(c_i) = F_Z(c_i)$ ($i = 1, 2$) と $X_n \rightarrow a$ in prob. より、 $N \in \mathbf{N}$ を

$$n \geq N \implies |F_{Z_n}(c_i) - F_Z(c_i)| < \frac{\varepsilon}{4} \quad (i = 1, 2), \quad P(A_n^c) < \frac{\varepsilon}{4}$$

となるように選ぶ。ただし、 $A_n = \left\{ |X_n - a| < \frac{a}{|c|+1} \delta' \right\}$, $\delta' = \min\{c - c_1, c_2 - c, 1\}$ とした。ここで、 A_n 上で $c_1 < c - \delta' < \frac{c}{a} X_n < c + \delta' < c_2$ かつ $X_n > 0$ となることに注意する。よって、

$$\begin{aligned} P\left(\frac{Z_n}{X_n} \leq \frac{c}{a}\right) &\leq P\left(\left\{\frac{Z_n}{X_n} \leq \frac{c}{a}\right\} \cap A_n\right) + P(A_n^c) = P\left(\left\{Z_n \leq \frac{c}{a} X_n\right\} \cap A_n\right) + P(A_n^c) \\ &\leq P(Z_n \leq c + \delta') + P(A_n^c) \leq P(Z_n \leq c_2) + \frac{\varepsilon}{4} \leq P(Z \leq c_2) + \frac{\varepsilon}{2}, \\ &< P(Z \leq c) + \varepsilon, \\ P\left(\frac{Z_n}{X_n} \leq \frac{c}{a}\right) &\geq P\left(\left\{\frac{Z_n}{X_n} \leq \frac{c}{a}\right\} \cap A_n\right) \geq P\left(\left\{Z_n \leq \frac{c}{a} X_n\right\} \cap A_n\right) \geq P(\{Z_n \leq c - \delta'\} \cap A_n) \\ &\geq P(Z_n \leq c - \delta') - P(A_n^c) \geq P(Z_n \leq c_1) - \frac{\varepsilon}{4} \geq P(Z \leq c_1) - \frac{\varepsilon}{2} \\ &> P(Z \leq c) - \varepsilon. \end{aligned}$$

以上より $n \geq N$ ならば $|P(Z_n/X_n \leq c/a) - F_Z(c)| < \varepsilon$ となるので主張を得る。 \square

命題 5.6 $X_n \rightarrow 0$ in prob. であり、確率変数の族 $\{Y_n\}$ は tight, すなわち、 $\forall \varepsilon > 0$ に対して $M > 0$ があって $\inf_n P(|Y_n| \leq M) \geq 1 - \varepsilon$ とすると、 $X_n Y_n \rightarrow 0$ in prob. となる。

証明: $\varepsilon > 0$ を任意にとっておく。条件より、 $M > 0$ を $P(|Y_n| > M) < \frac{\varepsilon}{2}$ とでき、さらに $\forall \delta > 0$ に対して $N \in \mathbf{N}$ を $n \geq N$ ならば $P(|X_n| \geq \frac{\delta}{M}) < \frac{\varepsilon}{2}$ とできる。このとき、 $n \geq N$ なら

$$\begin{aligned} P(|X_n Y_n| \geq \delta) &= P(|X_n Y_n| \geq \delta, |Y_n| \leq M) + P(|Y_n| > M) \\ &\leq P(M|X_n| \geq \delta, |Y_n| \leq M) + P(|Y_n| > M) \\ &\leq P(|X_n| \geq \frac{\delta}{M}) + P(|Y_n| > M) < \varepsilon. \quad \square \end{aligned}$$

以下、(A1)–(A3) に加え次を仮定する。 X は X_i と同じ分布をもつ確率変数とする。

(A4) $f(x|\theta)$ は θ について C^3 級。

(A5) $E_{\theta_0} \left[\frac{\partial}{\partial \theta} \log f(X|\theta_0) \right] = 0$, $I(\theta_0) := E_{\theta_0} \left[-\frac{\partial^2}{\partial \theta^2} \log f(X|\theta_0) \right] = E_{\theta_0} \left[\left\{ \frac{\partial}{\partial \theta} \log f(X|\theta_0) \right\}^2 \right] \in (0, \infty)$.

(A6) ある $c > 0$ と $M(x)$ が存在して、 $x \in A$ (cf. (A2)) と $|\theta - \theta_0| < c$ に対して $\left| \frac{\partial^3}{\partial \theta^3} \log f(x|\theta) \right| \leq M(x)$ であり $E_{\theta_0}[M(X)] < \infty$.

もし $\int f(x|\theta) dx$ が θ について 2 回微分可能なら (A5) は成立すると期待できる。実際、 $\int_{-\infty}^{\infty} f(x|\theta) dx = 1$ より

$$\int_{-\infty}^{\infty} \frac{\partial}{\partial \theta} f(x|\theta) dx = \frac{\partial}{\partial \theta} \left(\int_{-\infty}^{\infty} f(x|\theta) dx \right) = 0, \quad \text{同様に} \quad \int_{-\infty}^{\infty} \frac{\partial^2}{\partial \theta^2} f(x|\theta) dx = 0. \quad (5.2)$$

よって、 $\frac{\partial}{\partial \theta} \log f(x|\theta) = \frac{1}{f(x|\theta)} \frac{\partial f}{\partial \theta}(x|\theta)$, $\frac{\partial^2}{\partial \theta^2} \log f(x|\theta) = \frac{1}{f(x|\theta)} \frac{\partial^2 f}{\partial \theta^2}(x|\theta) - \left(\frac{1}{f(x|\theta)} \frac{\partial f}{\partial \theta}(x|\theta) \right)^2$ より、

$$E_{\theta_0} \left[\frac{\partial}{\partial \theta} \log f(x|\theta_0) \right] = E_{\theta_0} \left[\frac{1}{f(X|\theta_0)} \frac{\partial f}{\partial \theta}(X|\theta_0) \right] = \int_{-\infty}^{\infty} \frac{\partial}{\partial \theta} f(x|\theta_0) dx = 0, \quad (5.3)$$

$$\begin{aligned} E_{\theta_0} \left[\frac{\partial^2}{\partial \theta^2} \log f(x|\theta_0) \right] &= E_{\theta_0} \left[\frac{1}{f(X|\theta_0)} \frac{\partial^2 f}{\partial \theta^2}(X|\theta_0) \right] - E_{\theta_0} \left[\left\{ \frac{1}{f(X|\theta_0)} \frac{\partial f}{\partial \theta}(X|\theta_0) \right\}^2 \right] \\ &= \int_{-\infty}^{\infty} \frac{\partial^2}{\partial \theta^2} f(x|\theta_0) dx - E_{\theta_0} \left[\left\{ \frac{1}{f(X|\theta_0)} \frac{\partial f}{\partial \theta}(X|\theta_0) \right\}^2 \right] \\ &= -E_{\theta_0} \left[\left\{ \frac{\partial}{\partial \theta} \log f(X|\theta_0) \right\}^2 \right] = -I(\theta_0) \end{aligned} \quad (5.4)$$

となる。ここで、 $I(\theta)$ は Fisher 情報量 (前期の Cramér-Rao の定理のところ学んだ) であった。

定理 5.7 (最尤推定量の漸近有効性) X_1, X_2, \dots は i.i.d. で (A1)–(A6) の条件を満たすとする。尤度方程式 (5.1) の根 $\hat{\theta}_n = \hat{\theta}_n(x_1, \dots, x_n)$ は

$$\sqrt{n}(\hat{\theta}_n(\mathbf{X}) - \theta_0) \rightarrow N(0, 1/I(\theta_0)) \quad \text{in law} \quad (5.5)$$

を満たす。特に、最尤方程式 (5.1) がただ一つの解 $\hat{\theta}_n$ をもつ、すなわち、 $\hat{\theta}_n$ が最尤推定量であれば、(5.5) が成り立つ。ただし $\hat{\theta}_n(\mathbf{X}) = \hat{\theta}_n(X_1, \dots, X_n)$ とした。

証明: $\mathbf{x} = (x_1, \dots, x_n)$ を固定し、 $l'(\hat{\theta}_n|\mathbf{x})$ について θ_0 のまわりで Taylor の定理を適用すると

$$l'(\hat{\theta}_n|\mathbf{x}) = l'(\theta_0|\mathbf{x}) + (\hat{\theta}_n - \theta_0)l''(\theta_0|\mathbf{x}) + \frac{1}{2}(\hat{\theta}_n - \theta_0)^2 l'''(\theta_n^*|\mathbf{x})$$

となる。ここで、 $\theta_n^* = \theta_0 + h(\hat{\theta}_n - \theta_0)$, $0 < h < 1$, である。仮定より (左辺)=0 なので、

$$\sqrt{n}(\hat{\theta}_n(\mathbf{X}) - \theta_0) = \frac{\frac{1}{\sqrt{n}}l'(\theta_0|\mathbf{X})}{-\frac{1}{n}l''(\theta_0|\mathbf{X}) - \frac{1}{2n}(\hat{\theta}_n(\mathbf{X}) - \theta_0)l'''(\theta_n^*|\mathbf{X})}.$$

ここで、(A5) より中心極限定理から

$$\frac{1}{\sqrt{n}}l'(\theta_0|\mathbf{X}) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\partial}{\partial \theta} \log f(X_i|\theta_0) \rightarrow N(0, I(\theta_0)) \quad \text{in law.}$$

また、(A5) と大数の法則より、

$$\frac{1}{n}l''(\theta_0|\mathbf{X}) = \frac{1}{n} \sum_{i=1}^n \frac{\partial^2}{\partial \theta^2} \log f(X_i|\theta_0) \rightarrow E_{\theta_0} \left[\frac{\partial^2}{\partial \theta^2} \log f(X|\theta_0) \right] = -I(\theta_0) \quad \text{in prob.}$$

次に定理 5.4 より $\hat{\theta}_n \rightarrow \theta_0$ in prob. なので、 $\forall \varepsilon > 0$ に対し N を $n \geq N$ ならば $P_{\theta_0}(|\hat{\theta}_n - \theta_0| \geq \varepsilon) < \varepsilon/2$ とすると、 $|\theta_n^* - \theta_0| \leq |\hat{\theta}_n - \theta_0|$ より (A6) と大数の法則より

$$\left| \frac{1}{n}l'''(\theta_n^*|\mathbf{X})1_{\{|\hat{\theta}_n - \theta_0| < c\}} \right| \leq \frac{1}{n} \sum_{i=1}^n \left| \frac{\partial^3}{\partial \theta^3} \log f(X_i|\theta_n^*) \right| 1_{\{|\hat{\theta}_n - \theta_0| < c\}} \leq \frac{1}{n} \sum_{i=1}^n M(X_i) \rightarrow E_{\theta_0}[M(X)] \quad \text{in prob.}$$

よって、 $M_0 = E_{\theta_0}[M(X)]$ として、十分大きい n に対して

$$P_{\theta_0} \left(\left| \frac{1}{n}l'''(\theta_n^*|\mathbf{X}) \right| \geq M_0 + 1 \right) \leq P_{\theta_0}(|\hat{\theta}_n - \theta_0| \geq c) + P_{\theta_0} \left(\left| \frac{1}{n}l'''(\theta_n^*|\mathbf{X})1_{\{|\hat{\theta}_n - \theta_0| < c\}} \right| \geq M_0 + 1 \right) < \varepsilon$$

とできるので、命題 5.5, 5.6 より、 $Z \sim N(0, I(\theta_0))$ として

$$\sqrt{n}(\hat{\theta}_n(\mathbf{X}) - \theta_0) \rightarrow \frac{1}{I(\theta_0)}Z \quad \text{in law}$$

となるが、 $\frac{1}{I(\theta_0)}Z \sim N\left(0, \frac{1}{I(\theta_0)}\right)$ より結論を得る。□

この定理は最尤推定量 $\hat{\theta}_n(\mathbf{X})$ が漸近的に平均 θ_0 , 分散 $\frac{1}{n}I(\theta_0)^{-1}$ の正規分布に従うことを示している。

例 5.1 X_1, X_2, \dots が i.i.d. で各 X_i は指数分布 $\text{Ex}(\lambda)$, $\lambda > 0$, に従うとき、 λ の最尤推定量 $\hat{\lambda}$ とその漸近分布を求めよ。

解: $f(x|\lambda) = \lambda e^{-\lambda x}$ とする。尤度方程式

$$l'_n(\lambda|\mathbf{x}) = \frac{\partial}{\partial \lambda} \sum_{i=1}^n \log f(x_i|\lambda) = \frac{n}{\lambda} - (x_1 + \dots + x_n) = 0$$

より、最尤推定量は $\hat{\lambda}_n = \frac{1}{\bar{X}_n}$, ただし $\bar{X}_n = \frac{1}{n}(X_1 + \dots + X_n)$. Fisher 情報量は

$$I(\lambda) = -E\left[\frac{\partial^2}{\partial \lambda^2} \log f(X|\lambda)\right] = \frac{1}{\lambda^2}.$$

よって、 $\hat{\lambda}$ は漸近的に $N\left(\lambda, \frac{\lambda^2}{n}\right)$ に従う。 □

例 5.2 ある保険契約の各契約者の 1 年間のクレーム件数の分布は、互いに独立に同一の幾何分布 $Ge(p)$, $0 < p < 1$, に従うものとする。ある年度のクレーム件数を調べたところ、全契約者 10,000 人のクレーム件数は 500 件であった。このとき、この母集団の母数 p の 95% 信頼区間を、最尤推定量の漸近有効性を利用して求めよ。^{*4}

解: $f(x|p) = (1-p)^x p$, $x = 0, 1, 2, \dots$, とする。尤度方程式

$$l'_n(p|\mathbf{x}) = \frac{\partial}{\partial p} \sum_{i=1}^n \log f(x_i|p) = -\frac{x_1 + \dots + x_n}{1-p} + \frac{n}{p} = 0$$

より、最尤推定量は $\hat{p}_n = \frac{1}{1 + \bar{X}_n}$, ただし $\bar{X}_n = \frac{1}{n}(X_1 + \dots + X_n)$. Fisher 情報量は

$$I(p) = -E\left[\frac{\partial^2}{\partial p^2} \log f(X|p)\right] = E\left[\frac{X}{(1-p)^2} + \frac{1}{p^2}\right] = \frac{1}{(1-p)^2} \frac{1-p}{p} + \frac{1}{p^2} = \frac{1}{(1-p)p^2}.$$

よって、 $\hat{p} = \frac{1}{1 + \bar{X}}$ は漸近的に $N\left(p, \frac{(1-p)p^2}{n}\right)$ に従う。ここで、漸近分散 $\frac{(1-p)p^2}{n}$ は $\frac{(1-\hat{p})\hat{p}^2}{n}$ に十分近いと考え、実現値を代入して $\hat{p} = 1/(1 + \frac{500}{10000}) = 0.95238 \dots$ で

$$\hat{p} \pm 1.960 \sqrt{\frac{(1-\hat{p})\hat{p}^2}{n}} = 0.95238 \pm 1.960 \sqrt{\frac{0.04762 \cdot 0.95238^2}{10000}} = 0.95238 \pm 0.00407 = \begin{cases} 0.95645 \\ 0.94831 \end{cases}$$

となるので、95% 信頼区間は $0.9483 \leq p \leq 0.9565$ を得る。 □

問題 5.1 $a > 0$ を定数、 $\theta > 0$ を未知母数とする。 X_1, X_2, \dots が i.i.d. で X_i の密度関数を $f(x|\theta) = a\theta^a x^{a-1} e^{-(\theta x)^a}$, $x > 0$, とする。このとき、最尤推定量 $\hat{\theta}_n$ と、 $\hat{\theta}_n$ の漸近分布を求めよ。

未知母数が k 次元のとき、未知母数 $\boldsymbol{\theta} = (\theta_1, \dots, \theta_k)$ に対し、Fisher 情報量 $I(\boldsymbol{\theta})$ を次で定義する。

$$I(\boldsymbol{\theta}) = \begin{pmatrix} I_{11}(\boldsymbol{\theta}) & \dots & I_{1k}(\boldsymbol{\theta}) \\ \vdots & & \vdots \\ I_{k1}(\boldsymbol{\theta}) & \dots & I_{kk}(\boldsymbol{\theta}) \end{pmatrix},$$

$$I_{ij}(\boldsymbol{\theta}) = E_{\boldsymbol{\theta}} \left[\left\{ \frac{\partial}{\partial \theta_i} \log f(X|\boldsymbol{\theta}) \right\} \left\{ \frac{\partial}{\partial \theta_j} \log f(X|\boldsymbol{\theta}) \right\} \right] = -E_{\boldsymbol{\theta}} \left[\frac{\partial^2}{\partial \theta_i \partial \theta_j} \log f(X|\boldsymbol{\theta}) \right], \quad i, j = 1, \dots, k.$$

上式の変形は (5.4) と同様に示せる。

(A1)–(A6) と同様の条件のもと定理 5.8 が成立する。証明は省略する。ただし、(A5) では $I(\theta_0) \in (0, \infty)$ のかわりに、 $I(\boldsymbol{\theta})$ が正定値行列であることを仮定する。

定理 5.8 (最尤推定量の分散 (損保数理 p.2–2)) X_1, X_2, \dots は i.i.d. で上記の条件を満たすとする。尤度方程式

$$\frac{\partial}{\partial \theta_j} l_n(\boldsymbol{\theta}|\mathbf{x}) = \sum_{i=1}^n \frac{\partial}{\partial \theta_j} \log f(x_i|\boldsymbol{\theta}) = 0, \quad j = 1, \dots, k \quad (5.6)$$

^{*4} [IK] 岩沢宏和 黒田耕嗣 著 損害保険数理 (アクチュアリー数学シリーズ 4), 日本評論社, 2015

の根 $\widehat{\boldsymbol{\theta}}_n = \widehat{\boldsymbol{\theta}}_n(x_1, \dots, x_n)$ は

$$\sqrt{n}(\widehat{\boldsymbol{\theta}}_n(\mathbf{X}) - \boldsymbol{\theta}_0) \rightarrow N(0, I(\boldsymbol{\theta}_0)^{-1}) \quad \text{in law} \quad (5.7)$$

を満たす。特に、最尤方程式 (5.6) がただ一つの解 $\widehat{\boldsymbol{\theta}}_n$ をもつ、すなわち、 $\widehat{\boldsymbol{\theta}}_n$ が最尤推定量であれば、(5.7) が成り立つ。

この定理は最尤推定量 $\widehat{\boldsymbol{\theta}}_n(\mathbf{X})$ が漸近的に平均ベクトル $\boldsymbol{\theta}_0$ 、共分散行列 $\frac{1}{n}I(\boldsymbol{\theta}_0)^{-1}$ の k 次元正規分布に従うことを示している。

例 5.3 $\alpha, \beta > 0$ を未知母数とする。 X_1, X_2, \dots が i.i.d. で各 X_i は分布関数が $F(x|\alpha, \beta) = 1 - \left(\frac{\beta}{x+\beta}\right)^\alpha$ ($x \geq 0$) で与えられる確率変数であるとする。このとき、 α, β の最尤推定量 $\widehat{\alpha}_n, \widehat{\beta}_n$ の漸近分布を求めよ。

解: 密度関数は $f(x|\alpha, \beta) = \frac{\alpha\beta^\alpha}{(x+\beta)^{\alpha+1}}$ 。よって

$$\begin{aligned} \frac{\partial}{\partial \alpha} \log f(x|\alpha, \beta) &= \frac{1}{\alpha} + \log \beta - \log(x+\beta), & \frac{\partial}{\partial \beta} \log f(x|\alpha, \beta) &= \frac{\alpha}{\beta} - \frac{\alpha+1}{x+\beta}, \\ \frac{\partial^2}{\partial \alpha^2} \log f(x|\alpha, \beta) &= -\frac{1}{\alpha^2}, & \frac{\partial^2}{\partial \alpha \partial \beta} \log f(x|\alpha, \beta) &= \frac{1}{\beta} - \frac{1}{x+\beta}, & \frac{\partial^2}{\partial \beta^2} \log f(x|\alpha, \beta) &= -\frac{\alpha}{\beta^2} + \frac{\alpha+1}{(x+\beta)^2}. \end{aligned}$$

ここで、

$$E[(X+\beta)^{-k}] = \int_0^\infty \frac{\alpha\beta^\alpha}{(x+\beta)^{\alpha+1+k}} dx = \frac{\alpha}{(\alpha+k)\beta^k}, \quad k > -\alpha,$$

より、Fisher 情報量は $I(\alpha, \beta) = \begin{pmatrix} \frac{1}{\alpha^2} & -\frac{1}{(\alpha+1)\beta} \\ -\frac{1}{(\alpha+1)\beta} & \frac{1}{(\alpha+2)\beta^2} \end{pmatrix}$ 。よって、 $I(\alpha, \beta)^{-1}$ を計算して、 $(\widehat{\alpha}_n, \widehat{\beta}_n)$ は漸近的に $N\left(\begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \frac{1}{n} \begin{pmatrix} \alpha^2(\alpha+1)^2 & \alpha(\alpha+1)(\alpha+2)\beta \\ \alpha(\alpha+1)(\alpha+2)\beta & \frac{(\alpha+1)^2(\alpha+2)\beta^2}{\alpha} \end{pmatrix}\right)$ に従う。 \square

問題 5.2 X_1, X_2, \dots が i.i.d. で μ, σ^2 を未知母数とする対数正規分布に従う、すなわち、その密度関数が $f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2x}} \exp\left\{-\frac{(\log x - \mu)^2}{2\sigma^2}\right\}$, $x > 0$, であるとする。このとき、 μ, σ^2 の最尤推定量 $\widehat{\mu}_n, \widehat{\sigma}_n^2$ の漸近分布を求めよ。

定理 5.9 (最尤推定量の関数の分散 (損保数理 p.2-4)) k 次元確率変数の列 $\mathbf{X}_n = (X_{1n}, \dots, X_{kn})$ が漸的に平均 $\boldsymbol{\theta} = (\theta_1, \dots, \theta_k)$ 、共分散行列 $\frac{1}{n}\Sigma$ の正規分布に従い、 $\boldsymbol{\theta}$ と Σ はいずれも n に依存しないものと仮定する。 g を全微分可能な k 変数関数であると、 $G_n = g(X_{1n}, \dots, X_{kn})$ とすると、 G_n は漸的に平均 $g(\boldsymbol{\theta})$ 、分散 $\frac{1}{n}(\partial g)' \Sigma (\partial g)$ の正規分布に従う。ここで、 $(\partial g) = \begin{pmatrix} \frac{\partial g}{\partial \theta_1}(\boldsymbol{\theta}) & \dots & \frac{\partial g}{\partial \theta_k}(\boldsymbol{\theta}) \end{pmatrix}$ (縦ベクトル) である。

この定理は、定理 5.8 の仮定の下、最尤推定量 $\widehat{\boldsymbol{\theta}}_n$ の関数 $g(\widehat{\boldsymbol{\theta}}_n)$ は漸的に、平均 $g(\boldsymbol{\theta}_0)$ 、分散 $\frac{1}{n}(\partial g)' I(\boldsymbol{\theta}_0)^{-1} (\partial g)$ の正規分布に従うことを示している。

略証: 簡単のため $k = 2$ として示す。まず a, b が定数で、 (X, Y) が $N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}\right)$ に従うとき、 $aX + bY$ は平均 $a\mu_1 + b\mu_2$ で次の分散の正規分布に従うことに注意する:

$$V(aX + bY) = a^2V(X) + 2ab \text{Cov}(X, Y) + b^2V(Y) = \begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}. \quad (5.8)$$

$g(x_1, x_2)$ は全微分可能なので、多変数の平均値の定理より

$$G_n - g(\theta_1, \theta_2) = \frac{\partial g}{\partial \theta_1}(X_{1n}^*, X_{2n}^*)(X_{1n} - \theta_1) + \frac{\partial g}{\partial \theta_2}(X_{1n}^*, X_{2n}^*)(X_{2n} - \theta_2) \quad (5.9)$$

となる。ここで、 $(X_{1n}^*, X_{2n}^*) = (\theta_1, \theta_2) + h(X_{1n} - \theta_1, X_{2n} - \theta_2)$, $0 < h < 1$, である。仮定より

$$X_{1n} \rightarrow \theta_1 \text{ in prob.} \quad \text{かつ} \quad X_{2n} \rightarrow \theta_2 \text{ in prob.}$$

が証明できるので、

$$\frac{\partial g}{\partial \theta_1}(X_{1n}^*, X_{2n}^*) \rightarrow \frac{\partial g}{\partial \theta_1}(\theta_1, \theta_2) \text{ in prob.} \quad \text{かつ} \quad \frac{\partial g}{\partial \theta_2}(X_{1n}^*, X_{2n}^*) \rightarrow \frac{\partial g}{\partial \theta_2}(\theta_1, \theta_2) \text{ in prob.}$$

を得る。さらに、仮定より $(\sqrt{n}(X_{1n} - \theta_1), \sqrt{n}(X_{2n} - \theta_2))$ は平均 $(0, 0)$, 共分散行列 Σ の 2 次元正規分布に法則収束するので、その極限分布に従う確率変数を (Y_1, Y_2) とすると、(5.9) の両辺に \sqrt{n} を掛け $n \rightarrow \infty$ とすると、

$$\sqrt{n}(G_n - g(\theta_1, \theta_2)) \rightarrow \frac{\partial g}{\partial \theta_1}(\theta_1, \theta_2)Y_1 + \frac{\partial g}{\partial \theta_2}(\theta_1, \theta_2)Y_2 \text{ in law}$$

となるので、(5.8) に注意して主張を得る。 \square

例 5.4 例 5.2 で、この母集団の母平均 $\mu = \frac{1-p}{p}$ の 95% 信頼区間を、最尤推定量 \hat{p}_n の漸近有効性を利用して求めよ。

解: 例 5.2 より、最尤推定量 $\hat{p}_n = \frac{1}{1+\bar{X}}$ で、 $\sqrt{n}(\hat{p}_n - p) \rightarrow N(0, (1-p)p^2)$ であった。母平均を \hat{p}_n を用いて表すと $g(p) = \frac{1-p}{p}$ とするとき、 $g(\hat{p}_n)$ と表せるから、 $g(\hat{p}_n)$ は漸近的に平均 $g(p)$, 分散

$$\frac{1}{n}g'(p)(1-p)p^2g'(p) = \frac{1}{n}\left(-\frac{1}{p^2}\right)^2(1-p)p^2 = \frac{1-p}{np^2}$$

の正規分布に従うと考えられる。従って、95% 信頼区間は

$$g(\hat{p}_n) \pm u(0.025)\sqrt{\frac{1-\hat{p}_n}{n\hat{p}_n^2}} = \bar{x} \pm u(0.025)\sqrt{\frac{\bar{x}(1+\bar{x})}{n}} = 0.05 \pm 1.960\sqrt{\frac{0.05(1+0.05)}{10000}} = 0.05 \pm 0.00449$$

となるので、95% 信頼区間は $0.0455 \leq \mu \leq 0.0545$ を得る。 \square

問題 5.3 問題 5.2 の対数正規分布に従う母集団で、標本数を n とするとき、 $(\hat{\mu}_n, \hat{\sigma}_n^2)$ の漸近有効性を用いて、次の問いに答えよ。

(1) $\sigma = \sqrt{\hat{\sigma}_n^2}$ の漸近分布を求めよ。

(2) 母集団分布の平均の漸近分布を求めよ。ヒント: 対数正規分布での平均は $e^{\hat{\mu}_n + \frac{1}{2}\hat{\sigma}_n^2}$ であった。

5.2 極値問題

ここでは、損保数理の重要なテーマである極値問題を扱う。基本的な参考図書として次がある。

[R] S.I. Resnick: Extreme Values, Regular Variation and Point Processes, Springer, 1987

[TS] 高橋 倫也, 志村 隆彰: 極値統計学 (ISM シリーズ: 進化する統計数理), 近代科学社, 2016

定理 5.10 (型収束定理 (損保数理 p.10-55)) 確率変数列 $\{X_n\}$ が、定数 $a_n > 0, \alpha_n > 0, b_n, \beta_n \in \mathbf{R}$ と a.s. には定数でない確率変数 X, Y があって、

$$\frac{X_n - b_n}{a_n} \rightarrow X \text{ in law}, \quad \frac{X_n - \beta_n}{\alpha_n} \rightarrow Y \text{ in law}, \quad (5.10)$$

であるための必要十分条件は、

$$\lim_{n \rightarrow \infty} \frac{a_n}{\alpha_n} = a > 0, \quad \lim_{n \rightarrow \infty} \frac{b_n - \beta_n}{\alpha_n} = b \quad (5.11)$$

となることである。このとき、 a, b は一意的に定まり、 Y と $aX + b$ は同じ分布に従う。

ここで、確率変数 X が a.s. には定数ではないというのは、 $P(X = c) = 1$ となる定数 c が存在しないという
意味である。また、ある定数 $a > 0$ と $b \in \mathbf{R}$ があって、 $aX + b$ と Y が同じ分布に従うとき、すなわち、

$$F_Y(x) = P(aX + b \leq x) = F_X\left(\frac{x - b}{a}\right)$$

となるとき、 X と Y の分布は同じ型に属するという。

証明は難しいので省略する (cf. [R] pp.7-8)。直感的には、(5.10) より

$$Y \asymp \frac{X_n - \beta_n}{\alpha_n} = \frac{a_n}{\alpha_n} \frac{X_n - b_n}{a_n} + \frac{b_n - \beta_n}{\alpha_n} \asymp \frac{a_n}{\alpha_n} X + \frac{b_n - \beta_n}{\alpha_n}$$

であるが、 X, Y がともに a.s. には定数ではないことより (5.11) と必要十分となると理解できる。

定理 5.11 (Fisher-Tippett(損保数理 p.10-5)) X_1, X_2, \dots を i.i.d. とし、その最大値 $M_n = \max\{X_1, \dots, X_n\}$ を考える。このとき、a.s. には定数でない確率変数 Y と定数 $c_n > 0, d_n \in \mathbf{R}$ が存在して、

$$\frac{M_n - d_n}{c_n} \rightarrow Y \quad \text{in law} \quad (5.12)$$

とすれば、 Y の分布は次のいずれかの分布関数のもつ分布と同じ型に属する。

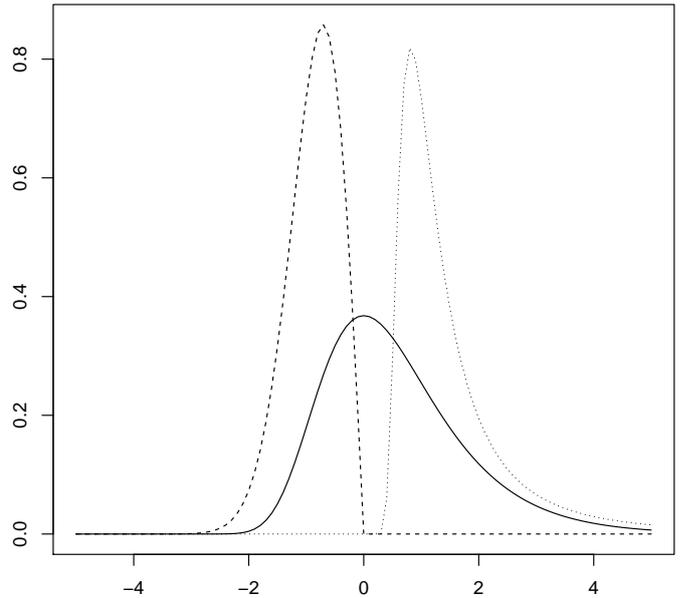
(Fréchet 分布) $\Phi_\alpha(x) = \begin{cases} \exp(-x^{-\alpha}), & x > 0, \\ 0, & x \leq 0. \end{cases}$

(Weibull 分布) $\Psi_\alpha(x) = \begin{cases} \exp(-(-x)^\alpha), & x \leq 0, \\ 1, & x > 0 \end{cases}$

(Gumbel 分布) $\Lambda(x) = \exp(-e^{-x})$.

ここで $\alpha > 0$ は定数である。

定理の $\Phi_\alpha(x), \Psi_\alpha(x), \Lambda(x)$ で与えられ分布を総称して極値分布という。その分布の密度関数のグラフは右図のようになる。($\alpha = 2$ とした。) どのグラフがどの分布に対応するかは、密度関数が正となる範囲を考えればすぐわかるであろう。



定理 5.11 の証明: X_i, Y の分布関数を

$$F(x) = P(X_i \leq x), \quad H(x) = P(Y \leq x)$$

とすると、 M_n の分布関数は $(F(x))^n$ であり、(5.12) より $H(x)$ の任意の連続点 $x \in \mathbf{R}$ に対して

$$P\left(\frac{M_n - d_n}{c_n} \leq x\right) = (F(c_n x + d_n))^n \rightarrow H(x), \quad n \rightarrow \infty. \quad (5.13)$$

(5.13) は n を $[nt]$ に置き換えても成り立つから、

$$P\left(\frac{M_{[nt]} - d_{[nt]}}{c_{[nt]}} \leq x\right) = (F(c_{[nt]} x + d_{[nt]}))^{[nt]} \rightarrow H(x), \quad n \rightarrow \infty. \quad (5.14)$$

ここで $[a]$ は a の整数部分を表すものとする。同様に、 $M_{[nt]}$ の分布関数は $(F(x))^{[nt]}$ であるから、

$$P\left(\frac{M_{[nt]} - d_n}{c_n} \leq x\right) = (F(c_n x + d_n))^{[nt]} = \left\{ (F(c_n x + d_n))^n \right\}^{[nt]/n} \rightarrow H(x)^t \quad (5.15)$$

を得る。(5.14), (5.15) より、 $\{H(x)\}^t$ を分布関数にもつ、したがって a.s. に定数ではない確率変数 Y_t があって、

$$\frac{M_{[nt]} - d_{[nt]}}{c_{[nt]}} \rightarrow Y \quad \text{in law}, \quad \frac{M_{[nt]} - d_n}{c_n} \rightarrow Y_t \quad \text{in law},$$

となる。よって、型収束定理より、

$$\lim_{n \rightarrow \infty} \frac{c_n}{c_{[nt]}} = \gamma(t) > 0, \quad \lim_{n \rightarrow \infty} \frac{d_n - d_{[nt]}}{c_{[nt]}} = \delta(t) \quad (5.16)$$

となる $\gamma(t), \delta(t)$ ($t > 0$) が存在し、 $\gamma(t)Y_t + \delta(t)$ と Y が同じ分布に従う:

$$\{H(x)\}^t = P(Y_t \leq x) = P(Y \leq \gamma(t)x + \delta(t)) = H(\gamma(t)x + \delta(t)). \quad (5.17)$$

ここで、 t のかわりに st を代入して

$$\{H(x)\}^{st} = H(\gamma(st)x + \delta(st)).$$

一方、

$$\{H(x)\}^{st} = \{(H(x))^s\}^t = \{H(\gamma(s)x + \delta(s))\}^t = H(\gamma(t)\{\gamma(s)x + \delta(s)\} + \delta(t)).$$

上の二式が任意の x に対して成立するので、後で述べる補題 5.12 より次の等式を得る。

$$\gamma(st) = \gamma(s)\gamma(t), \quad \delta(st) = \gamma(t)\delta(s) + \delta(t), \quad s, t > 0 \quad (5.18)$$

次に、 $\gamma(e) = e^\theta$, $\theta \in \mathbf{R}$, とすると、(5.18) より $s, t > 0$ と $n \in \mathbf{N}$ に対して $\gamma(s^n) = \gamma(s)^n$ であるから、 $t = s^n$ として $\gamma(t^{1/n}) = \gamma(t)^{1/n}$ となり、したがって、 $\forall r \in \mathbf{Q}$ に対して $\gamma(e^r) = \gamma(e)^r = (e^\theta)^r = (e^\theta)^r$ を得る。ここで、(5.16) より $\gamma(t)$ は可測関数なので次を得る*5。

$$\gamma(t) = t^\theta, \quad t > 0. \quad (5.19)$$

(a) $\theta = 0$ のとき: $\gamma(t) \equiv 1$ なので (5.18) より $\delta(st) = \delta(s) + \delta(t)$. 従って、上と同様に $\delta(e) = c$ とすると、 $\forall r \in \mathbf{Q}$ に対して $\delta(e^r) = r\delta(e) = rc$ となるので、 $\delta(t)$ の可測性より

$$\delta(t) = \delta(e^{\log t}) = c \log t, \quad t > 0,$$

となり、(5.17) より

$$\{H(x)\}^t = H(x + c \log t), \quad t > 0. \quad (5.20)$$

もし $c = 0$ なら、 $H(x) = 0, 1$ となり、これは Y が a.s. に定数であることを意味するので、 $c \neq 0$ となる。次に、 $0 \leq H(x) \leq 1$ より $\{H(x)\}^t$ は t について非増加なので、 $c < 0$ を得る。

次にある $x_0 \in \mathbf{R}$ があって $H(x_0) = 1$ とすると、(5.20) より $H(x_0 + c \log t) = 1, \forall t > 0$. これは $\forall x \in \mathbf{R}$ に対し $H(x) = 1$ を意味しており、 $H(x)$ が分布関数であることに矛盾する。したがって $\forall x \in \mathbf{R}$ に対し $H(x) < 1$. 同様に、 $x_0 \in \mathbf{R}$ があって $H(x_0) = 0$ とすると、 $\forall x \in \mathbf{R}$ に対し $H(x) = 0$ となり矛盾する。したがって $\forall x \in \mathbf{R}$ に対し $H(x) > 0$. (5.20) に $x = 0$ を代入して、

$$\{H(0)\}^t = H(c \log t).$$

ここで、 $c' = -1/c > 0$, $d \in \mathbf{R}$ を $\exp\{-e^{-d}\} = H(0) \in (0, 1)$ とし $x = c \log t$ を代入すると、 $t > 0$ のとき $-\infty < x < \infty$ で

$$H(x) = \{H(0)\}^{e^{x/c}} = \exp\{-e^{-d}e^{-c'x}\} = \Lambda(c'x + d).$$

*5 証明はかなり面倒なので省略する。詳しくは、例えば次の [BGT] pp.4-5 を見よ。

[BGT] N.H. Bingham, C.M. Goldie and J.L. Teugels: Regular Variation, Cambridge University Press, 1987.

(b) $\theta > 0$ のとき: (5.18) の第 2 式より、

$$\gamma(t)\delta(s) + \delta(t) = \gamma(s)\delta(t) + \delta(s)$$

より、 $t \neq 1, s \neq 1$ のとき

$$\frac{\delta(s)}{1-\gamma(s)} = \frac{\delta(t)}{1-\gamma(t)},$$

すなわち、 $\frac{\delta(s)}{1-\gamma(s)}$ は定数関数なのでそれを c とする。よって、 $t \neq 1$ のとき

$$\delta(t) = \delta(s) \frac{1-\gamma(t)}{1-\gamma(s)} = c(1-t^\theta)$$

であり、

$$\{H(x)\}^t = H(t^\theta x + c(1-t^\theta)) = H(t^\theta(x-c) + c) \quad (5.21)$$

特に $x = c$ として $H(c)^t = H(c)$ となるので、 $H(c) = 0$ または 1 。もし $H(c) = 0$ なら、ある $x > c$ があって $H(x) > 0$ となるので、右連続性より

$$\{H(x)\}^t = H(t^\theta(x-c) + c) \rightarrow H(c) = 0 \quad \text{as } t \rightarrow +0.$$

一方、 $a = H(x) > 0$ なら $a^t \rightarrow 1$ as $t \rightarrow +0$ となるので矛盾する。これより $H(x) \equiv 0$ となり、これは $H(x)$ が分布関数であることに矛盾する。したがって、 $H(c) = 1$ となる。

次に、 $0 < H(c-1) < 1$ を示す。 $H(c-1) = 1$ とすると、(5.21) より $H(-t^\theta + c) = H(t^\theta(c-1-c) + c) = \{H(c-1)\}^t = 1, t > 0$, となるので、 $H \equiv 1$ となり H が分布関数であることに矛盾する。また、 $H(c-1) = 0$ とすると、全く同様に $H \equiv 0$ となり矛盾する。したがって、 $p > 0$ があって $H(c-1) = \exp\{-p^\alpha\}$ とできる。ただし、 $\alpha = 1/\theta$ とした。このとき、 $x < c$ に対して、 $x - c = -t^\theta$ とすると $t = (c-x)^\alpha$ に注意して (5.21) を用い

$$H(x) = H(x - c + c) = H(-t^\theta + c) = \{H(c-1)\}^t = \exp\{-(p(c-x))^\alpha\} = \Psi_\alpha(p(x-c)).$$

(c) $\theta < 0$ のとき: (b) のときと同様に c を定めると、(5.21) が成り立つ。よって、 $H(c) = 0, 1$ となるが、もし $H(c) = 1$ ならある $x < c$ があって $H(x) < 1$ となるが、

$$H(c-0) = \lim_{t \rightarrow \infty} H(t^\theta(x-c) + c) = \lim_{t \rightarrow \infty} \{H(x)\}^t = 0$$

となり、 $Y = c$ a.s. となり矛盾する。よって $H(c) = 0$ 。

さらに、(b) とまったく同様に $0 < H(1+c) < 1$ が示されるので、 $p > 0$ があって $H(1+c) = \exp\{-p^{-\alpha}\}$ とできる。ただし、 $\alpha = -1/\theta$ とした。このとき、 $x > c$ に対して、 $x - c = t^\theta$ とすると $t = (x-c)^{-\alpha}$ に注意して (5.21) を用い

$$H(x) = H(x - c + c) = H(t^\theta + c) = \{H(1+c)\}^t = \exp\{-(p(x-c))^{-\alpha}\} = \Phi_\alpha(p(x-c)). \quad \square$$

補題 5.12 X を a.s. には定数ではない確率変数とし $F(x)$ をその分布関数とする。このとき、定数 $a > 0, c > 0, b, d \in \mathbf{R}$ が、 $\forall x \in \mathbf{R}$ に対して $F(ax+b) = F(cx+d)$ を満たせば、 $a = c, b = d$ となる。

証明: $a \geq c$ として一般性を失わない。 $G(x) = F(cx+d)$ とすると、

$$F(ax+b) = F\left(c\left(\frac{a}{c}x + \frac{b-d}{c}\right) + d\right) = G\left(\frac{a}{c}x + \frac{b-d}{c}\right)$$

となることに注意し、 $\alpha = \frac{a}{c} \geq 1, \beta = \frac{b-d}{c}$ と表す。仮定より、 $G(x) = G(\alpha x + \beta)$ であるが、これを繰り返して代入して

$$G(x) = G(\alpha x + \beta) = G(\alpha(\alpha x + \beta) + \beta) = G(\alpha^2 x + (\alpha + 1)\beta)$$

$$= G(\alpha\{\alpha^2x + (\alpha + 1)\beta\} + \beta) = \cdots = G(\alpha^n x + (\alpha^{n-1} + \cdots + 1)\beta) \quad (5.22)$$

を得る。よって、 $\alpha > 1$ であれば

$$G(x) = G\left(\alpha^n x + \frac{\alpha^n - 1}{\alpha - 1}\beta\right) = G\left(\alpha^n \left(x + \frac{\beta}{\alpha - 1}\right) - \frac{\beta}{\alpha - 1}\right)$$

であるが、 $F(x)$ は a.s. には定数ではない確率変数の分布関数だから、 $G(x)$ もそうなるので、ある $0 < p < 1$ と $x_0 \neq -\frac{\beta}{\alpha - 1}$ があって、 $G(x_0) = p$ とできる。このとき、 $n \rightarrow \infty$ とすると

$$p = G(x_0) = G\left(\alpha^n \left(x_0 + \frac{\beta}{\alpha - 1}\right) - \frac{\beta}{\alpha - 1}\right) \rightarrow \begin{cases} 1, & x_0 + \frac{\beta}{\alpha - 1} > 0, \\ 0, & x_0 + \frac{\beta}{\alpha - 1} < 0, \end{cases}$$

となるので、いずれにせよ矛盾する。よって $\alpha = 1$ 。ここで、もし $\beta \neq 0$ なら、(5.22) で $n \rightarrow \infty$ とすると

$$p = G(x_0) = G(x_0 + n\beta) \rightarrow \begin{cases} 1, & \beta > 0, \\ 0, & \beta < 0, \end{cases}$$

となり矛盾する。したがって、 $\beta = 0$ となる。□

例 5.5 X_1, X_2, \dots が i.i.d. でその分布と定数 $c_n > 0$, $d_n \in \mathbf{R}$ が次の (1)–(3) で与えられるとき、定理 5.11 の極限分布 Y の分布関数を求め、どの分布と同じ型か調べよ ($\Phi_\alpha, \Psi_\alpha, \Lambda$ で表せ)。

(1) Parate 分布に従う: $f_X(x) = \begin{cases} \beta x^{-\beta-1}, & x \geq 1, \\ 0, & x < 1, \end{cases}$ で、 $c_n = n^{1/\beta}$, $d_n = 0$ 。ただし $\beta > 0$ は定数。

(2) ベータ分布 $\text{BETA}(1, b)$ に従う: $f_X(x) = \begin{cases} b(1-x)^{b-1}, & 0 < x < 1, \\ 0, & \text{その他}, \end{cases}$ で、 $c_n = n^{-1/b}$, $d_n = 1$ 。

(3) 指数分布 $\text{Ex}(\lambda)$ に従い、 $c_n = \frac{1}{\lambda}$, $d_n = \frac{1}{\lambda} \log n$ 。

解: $\Phi_\alpha(x), \Psi_\alpha(x), \Lambda(x)$ はすべて $-\infty < x < \infty$ で連続なので (5.13) より $\lim_{n \rightarrow \infty} \{F(c_n x + d_n)\}^n$ をそれで表せばよい。

(1) X_i の分布関数は $F(x) = \begin{cases} 1 - x^{-\beta}, & x \geq 1, \\ 0, & x < 1. \end{cases}$

$x \leq 0$ のとき $F(c_n + d_n) = 0$ 。 $x > 0$ のとき n が十分大なら $n^{1/\beta}x > 1$ となることに注意すると、

$$\{F(c_n x + d_n)\}^n = \{1 - (n^{1/\beta}x)^{-\beta}\}^n = \left(1 - \frac{x^{-\beta}}{n}\right)^n \rightarrow e^{-x^{-\beta}} = \Phi_\beta(x).$$

以上より、極限分布の分布関数は $\Phi_\beta(x)$ となる。(2), (3) は演習問題とする。□

問題 5.4 例 5.5 (2), (3) を解け。

例題 5.6 独立に平均 9 の指数分布に従う X_1, X_2, \dots, X_n の最大値を M_n とする。このとき、 M_{81} が 81 以上である確率を、以下の 2 通りの方法で求めよ。

(1) 最大値の分布関数から直接計算する方法。

(2) 最大値を正規化した分布を極値分布により近似 (cf. 例 5.5 (3)) して計算する方法。

解: (1) $x \geq 0$ に対して、 $F_{X_i}(x) = 1 - e^{-x/9}$ であるから、

$$P(M_{81} \geq 81) = 1 - P(M_{81} < 81) = 1 - (1 - e^{-81/9})^{81} = 0.00994700 \dots$$

(2) 例 5.5 (3) より $P\left(\frac{M_n - 9 \log n}{9} \leq x\right) \rightarrow e^{-e^{-x}}$, $n \rightarrow \infty$, となるから、

$$\begin{aligned} P(M_{81} \geq 81) &= 1 - P\left(\frac{M_{81} - 9 \log 81}{9} \leq \frac{81 - 9 \log 81}{9}\right) = 1 - \exp\{-e^{-9 + \log 81}\} \\ &= 1 - \exp\{-81e^{-9}\} = 0.00994639 \dots \quad \square \end{aligned}$$

問題 5.5 X_1, X_2, \dots, X_n は i.i.d. で例 5.5 (1) で $\beta = \frac{9}{7}$ とした Parate 分布に従うとし、その最大値を M_n とする。このとき、 M_{128} が 128 以上である確率を、以下の 2 通りの方法で求めよ。(ヒント: $128 = 2^7$.)

(1) 最大値の分布関数から直接計算する方法。

(2) 最大値を正規化した分布を極値分布により近似 (cf. 例 5.5 (1)) して計算する方法。ただし $e^{-1} = 0.36788$ として計算せよ。

定義 5.2 (一般化極値分布) $\sigma > 0, \mu \in \mathbf{R}$ のとき、分布関数が

$$H_{\xi, \mu, \sigma}(x) = \exp\left\{-\left(1 + \xi \frac{x - \mu}{\sigma}\right)^{-1/\xi}\right\}, \quad 1 + \xi \frac{x - \mu}{\sigma} > 0, \quad (5.23)$$

である分布のことを一般化極値分布あるいは **GEV 分布** (generalized extreme value distribution) という。ただし、 $\xi = 0$ のときは

$$H_{0, \mu, \sigma}(x) = \lim_{\xi \rightarrow 0} H_{\xi, \mu, \sigma}(x) = \exp\left\{-e^{-\frac{x - \mu}{\sigma}}\right\}, \quad -\infty < x < \infty, \quad (5.24)$$

とする。 ξ, μ, σ はそれぞれ形状パラメータ、位置パラメータ。尺度パラメータと呼ばれ、 $\mu = 0, \sigma = 1$ としたとき、特に $H_\xi(x)$ と記す。

GEV と $\Phi_\alpha(x), \Psi_\alpha(x), \Lambda(x)$ は次のように関係づけられる。

- Fréchet 分布 $\Phi_\alpha(x) = H_{1/\alpha}\left(\frac{x-1}{1/\alpha}\right), \quad \xi = 1/\alpha > 0$
- Gumbel 分布 $\Lambda(x) = H_0(x), \quad \xi = 0$
- Weibull 分布 $\Psi_\alpha(x) = H_{-1/\alpha}\left(\frac{x+1}{1/\alpha}\right), \quad \xi = -1/\alpha < 0$

命題 5.13 $\alpha > 0, Z \sim \text{Ex}(1)$ とすると、次が成立する。

(1) $X = -\log Z$ は Gumbel 分布 Λ に従い、 $M_X(t) = E[e^{tX}] = \Gamma(1-t)$ 。特に、 $E[X] = \gamma, V(X) = \pi^2/6$ 。ここで、 $\gamma = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{2} + \dots + \frac{1}{n} - \log n\right) = 0.577215664 \dots$ (Euler の定数) を表す。

(2) $X = Z^{-1/\alpha}$ は Fréchet 分布 Φ_α に従い、 $E[X^n] = \Gamma\left(1 - \frac{n}{\alpha}\right), 1 - \frac{n}{\alpha} > 0$ 。

(3) $X = -Z^{1/\alpha}$ は Weibull 分布 Ψ_α に従い、 $E[X^n] = (-1)^n \Gamma\left(1 + \frac{n}{\alpha}\right), 1 + \frac{n}{\alpha} > 0$ 。

証明: (1) のみ示し、残りは演習問題とする。 X の分布とその積率母関数については

$$P(X \leq x) = P(Z \geq e^{-x}) = \int_{e^{-x}}^{\infty} e^{-z} dz = e^{-e^{-x}} = \Lambda(x),$$

$$M_X(t) = E[e^{tX}] = E[Z^{-t}] = \int_0^{\infty} z^{-t} e^{-z} dz = \Gamma(1-t)$$

より従う。次に、 X のキミュラント母関数 $\psi_X(t) = \log M_X(t)$ に関して、 $\psi'_X(0) = E[X], \psi''_X(0) = V(X)$ となることに注意する。これを計算するため、ガンマ関数の Weierstrass の公式

$$\frac{1}{\Gamma(x)} = xe^{\gamma x} \prod_{p=1}^{\infty} \left(1 + \frac{x}{p}\right) e^{-x/p}$$

を用いる。(証明は特殊関数の教科書を調べてください。) これより、

$$\psi_X(t) = \log \Gamma(1-t) = -\log(1-t) - \gamma(1-t) - \sum_{p=1}^{\infty} \left\{ \log \left(1 + \frac{1-t}{p}\right) - \frac{1-t}{p} \right\},$$

$$\psi'_X(t) = \frac{1}{1-t} + \gamma - \sum_{p=1}^{\infty} \left\{ \frac{-1/p}{1 + (1-t)/p} + \frac{1}{p} \right\} = \frac{1}{1-t} + \gamma - \sum_{p=1}^{\infty} \left\{ \frac{1}{p} - \frac{1}{p+1-t} \right\},$$

$$\psi''_X(t) = \frac{1}{(1-t)^2} + \sum_{p=1}^{\infty} \frac{1}{(p+1-t)^2}.$$

以上より、

$$E[X] = \psi'_X(0) = 1 + \gamma - \sum_{p=1}^{\infty} \left\{ \frac{1}{p} - \frac{1}{p+1} \right\} = 1 + \gamma - 1 = \gamma,$$

$$V(X) = \psi''_X(0) = 1 + \sum_{p=1}^{\infty} \frac{1}{(p+1)^2} = \sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}. \quad \square$$

問題 5.6 命題 5.13 (2), (3) を示せ。

定義 5.3 (最大吸引域) 定理 5.11 で X_i を F , (5.12) を満たす Y の分布関数を H とするとき、定数 $c_n > 0$, $d_n \in \mathbf{R}$ が存在して、ある ξ に対して $H(x) = H_\xi(x)$ できる。このとき、 F は一般化極値分布 H_ξ の最大吸引域 (maximum domain of attraction) に属するといひ、 $F \in MDA(H_\xi)$ と表す。ただし、 H_ξ のかわりに Φ_α , Λ , Ψ_α を用いることも多い。

多くの確率分布が、上記のどれかの最大吸引域に属することが知られている。そのための必要十分条件や、もう少しチェックしやすい十分条件も知られている。ただし、連続型確率分布であっても一般化極値分布の最大吸引域に属さない例が知られているし、また、Poisson 分布や幾何分布も一般化極値分布の最大吸引域に属さないことが知られている。詳しくは [R], [TS] を参照のこと。

次の表は三つのタイプの極限分布に関するまとめである ([TS], pp.35-36 より)。ただし、 $\bar{F}(x) = 1 - F(x)$ は末尾分布関数、緩慢関数の定義は定理 5.16 の次にある。

• Fréchet 分布 の最大吸引域

Fréchet 分布	$\Phi_\alpha(x) = \exp(-x^{-\alpha}), \quad x > 0, \alpha > 0$
$F \in MDA(\Phi_\alpha)$	$\omega_F = \infty, \bar{F}(x) = x^{-\alpha}l(x), l(x)$ は緩慢関数
基準化定数	$c_n = F^{\leftarrow}(1 - 1/n), d_n = 0$
極限への収束	$M_n/c_n \rightarrow \Phi_\alpha$ in law
例: 裾の厚い (fat-tailed) 分布	
Cauchy 分布	$f(x) = 1/[\pi(1+x^2)], \quad x \in \mathbf{R}, c_n = n/\pi, \alpha = 1$
Parate 分布	$\bar{F}(x) \sim Kx^{-\alpha}, K, \alpha > 0, c_n = (Kn)^{1/\alpha}$
対数ガンマ分布	$f(x) = \frac{\alpha^\beta}{\Gamma(\beta)}(\log x)^{\beta-1}x^{-\alpha-1}, x > 1, \alpha, \beta > 0,$ $c_n = [(\log n)^{\beta-1}n/\Gamma(\beta)]^{1/\alpha}$

• Weibull 分布 の最大吸引域

Weibull 分布	$\Psi_\alpha(x) = \exp(-(-x)^\alpha), \quad x < 0, \alpha > 0$
$F \in MDA(\Psi_\alpha)$	$\omega_F < \infty, \bar{F}(\omega_F - x^{-1}) = x^{-\alpha}l(x), l(x)$ は緩慢関数
基準化定数	$c_n = \omega_F - F^{\leftarrow}(1 - 1/n), d_n = \omega_F$
極限への収束	$(M_n - \omega_F)/c_n \rightarrow \Psi_\alpha$ in law
例: 有限な上限をもつ分布	
一様分布	$f(x) = 1, \quad 0 < x < 1, c_n = 1/n, d_n = 1, \quad \alpha = 1$
ω_F でベキ	$\bar{F}(x) = K(\omega_F - x)^\alpha, \omega_F - K^{-1/\alpha} \leq x \leq \omega_F, K, \alpha > 0,$ $c_n = (Kn)^{-1/\alpha}, d_n = \omega_F$
ベータ分布 BETA(a, b)	$f(x) = \frac{1}{B(a,b)}x^{a-1}(1-x)^{b-1}, 0 < x < 1, a, b > 0,$ $c_n = \left[\frac{\Gamma(a)\Gamma(b+1)}{n\Gamma(a+b)} \right]^{1/b}, d_n = 1, \quad \alpha = b$

• Gumbel 分布 の最大吸引域

Gumbel 分布	$\Lambda(x) = \exp(-e^{-x}), x \in \mathbf{R}$
$F \in MDA(\Lambda)$	$\omega_F \leq \infty, \bar{F}(x) = c(x) \exp(-\int_{\alpha_F}^x \frac{g(t)}{a(t)} dt), \alpha_F < x < \omega_F,$ ただし、 $x \rightarrow \omega_F$ のとき $c(x) \rightarrow c > 0, g(x) \rightarrow 1, a'(x) \rightarrow 0.$
基準化定数	$d_n = F^{\leftarrow}(1 - 1/n), c_n = a(d_n)$
極限への収束	$(M_n - d_n)/c_n \rightarrow \Lambda$ in law
例: 統計学でよく用いられる分布	
ガンマ分布 $\Gamma(\lambda, 1)$	$f(x) = \frac{1}{\Gamma(\lambda)} x^{\lambda-1} e^{-x}, x > 0, \lambda > 0$ $c_n = 1, d_n = \log n + (\lambda - 1) \log \log n - \log \Gamma(\lambda)$
正規分布 $N(0, 1)$	$f(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2}), x \in \mathbf{R}$ $c_n = (2 \log n)^{-1/2}, d_n = \sqrt{2 \log n} - [\log \log n + \log(4\pi)] / (2\sqrt{2 \log n})$
対数正規分布 $LN(\mu, \sigma^2)$	$f(x) = \frac{1}{\sqrt{2\pi\sigma x}} \exp(-\frac{(\log x - \mu)^2}{2\sigma^2}), x > 0, \mu \in \mathbf{R}, \sigma > 0$ $c_n = \sigma(2 \log n)^{-1/2} d_n, d_n = \exp\{\mu + \sigma[\sqrt{2 \log n} - \frac{\log \log n + \log(4\pi)}{2\sqrt{2 \log n}}]\}$
逆ガウス分布	$f(x) = (\frac{\lambda}{2\pi x^3}) \exp[-\frac{\lambda(x-\mu)^2}{2\mu^2 x}], x > 0, \mu, \lambda > 0,$ $c_n = 2\mu^2/\lambda, d_n = \mu^2\{2 \log n - 3 \log \log n + \log \frac{\lambda^2}{4\pi\mu^2} + \frac{2\lambda}{\mu}\}/\lambda$
$\omega_F < \infty$ で 指数挙動	$\bar{F}(x) = K \exp(-\frac{a}{\omega_F - x}), x < \omega_F, a, K > 0,$ $c_n = a/[\log(Kn)]^2, d_n = \omega_F - a/\log(Kn)$

では、データが与えられたときそれがどの一般化極値分布の最大吸引域かが問題となるが、それには通常、最尤法が用いられる。具体的には、いくつかのブロックに分けて考えるブロック最大値モデル、ブロックの最大値だけでなく、ある閾値を超過したデータを対象とした閾値超過モデルがある。詳しくは損保数理の教科書、さらにはそこに出展されている次の書籍の第7章を参照されたい。

[MFE] A.J. McNeil, R. Frey, P. Embrechts: 定量的リスク管理 -基礎概念と数理技法-, 共立出版, 2008.

5.3 安定分布

極値問題では、独立同分布をもつ確率変数の最大値 M_n について、その「型」(cf. 定理 5.10) を調べた。この節ではその和 $S_n = X_1 + \dots + X_n$ について考える。

中心極限定理によると、もし $\sigma^2 := V(X_1) < \infty$ であれば、

$$\frac{1}{\sqrt{n}}(S_n - nE[X_1]) \rightarrow N(0, \sigma^2) \quad \text{in law}$$

となる。一方、各 X_k が Cauchy 分布に従う、すなわちその密度関数が $f(x) = \frac{1}{\pi(1+x^2)}$ であるとき、その特性関数は $\phi_{X_k}(t) = e^{-|t|}$ であったから、 $Y_n = \frac{1}{n}S_n$ に対し、

$$\phi_{Y_n}(t) = E[e^{itY_n}] = \prod_{k=1}^n E[e^{i\frac{t}{n}X_k}] = \prod_{k=1}^n \phi_{X_k}(t/n) = (e^{-|t/n|})^n = e^{-|t|}$$

となり、 Y_n も Cauchy 分布に従う、特に、「 $Y_n \rightarrow$ Cauchy 分布 in law」を得る。では、一般に S_n についてどんな「型」がありうるのだろうか。一般に次が知られている。

定義 5.4 (安定分布) 確率変数 X が安定分布に従うとは、任意の $k \in \mathbf{N}$ と X と同分布に従う独立な X_1, \dots, X_k に対して、ある定数 $a_k > 0, b_k$ があって、 $X_1 + \dots + X_k \sim a_k X + b_k$ となる時にいう。

定理 5.14 確率変数 Z が、ある定数 $a_n > 0, b_n$ に対して $\frac{S_n - b_n}{a_n} \rightarrow Z$ in law となるための必要十分条件は、 Z が安定分布に従うことである。

定理 5.15 確率変数 Z が安定分布に従うとき、 Z は正規分布であるか、指数 $\alpha \in (0, 2)$ と定数 β, m_1, m_2 があって、その特性関数 $\phi_Z(t)$ は次で与えられる。

$$\phi_Z(t) = \exp \left\{ i\beta t + m_1 \int_0^\infty \left(e^{itx} - 1 - \frac{itx}{1+x^2} \right) \frac{dx}{x^{1+\alpha}} + m_2 \int_{-\infty}^0 \left(e^{itx} - 1 - \frac{itx}{1+x^2} \right) \frac{dx}{|x|^{1+\alpha}} \right\} \quad (5.25)$$

さらに、指数 α が $0 < \alpha < 2$ であれば、 $c \in \mathbf{R}, d > 0$ と $|\theta| \leq 1$ を用いて

$$\phi_Z(t) = \begin{cases} \exp \{ ict - d|t|^\alpha \left(1 + i\theta \frac{t}{|t|} \tan \frac{\pi\alpha}{2} \right) \}, & \alpha \neq 1, \\ \exp \{ ict - d|t| \left(1 + i\theta \frac{t}{|t|} \frac{2}{\alpha} \log |t| \right) \}, & \alpha = 1, \end{cases}$$

と変形できる。

証明は、例えば次の第 9 章にある。

[B] L. Breiman, Probability, Addison-Wesley, 1968. (Classics in applied mathematics, 7, Society for Industrial and Applied Mathematics, 1992. Reprint 版)

ここでは、次の第 3 章 7 節に述べられている次の定理を紹介することとする。

[D] Durrett, R.: Probability Theory and Examples, 4th ed., Cambridge University Press, 2010.

定理 5.16 X_1, X_2, \dots は i.i.d. で、次を満たすとする。

(i) $\lim_{x \rightarrow \infty} \frac{P(X_1 > x)}{P(|X_1| > x)} = \theta \in [0, 1],$

(ii) $0 < \alpha < 2$ と緩慢関数 $L(x)$ があって、 $P(|X_1| > x) = x^{-\alpha} L(x).$

このとき、 $S_n = X_1 + \dots + X_n$ に対して、

$$a_n = \inf \{ x; P(|X_1| > x) \leq 1/n \}, \quad b_n = nE[X_1 1_{\{|X_1| \leq a_n\}}]$$

とすると、a.s. には定数でない確率変数 Z があって、 $\frac{S_n - b_n}{a_n} \rightarrow Z$ in law となる。ここで、 Z の特性関数は (5.32) あるいは (5.25) で $m_1 = \alpha\theta, m_2 = \alpha(1 - \theta), \beta$ は (5.33) として与えられる。

ここで、 $L(x)$ が緩慢関数であるとは、 $\forall t > 0$ に対して $\lim_{x \rightarrow \infty} \frac{L(tx)}{L(x)} = 1$ となるときにいう。例えば、 $\log x$ は緩慢関数である。一方、 x^α は $\alpha \neq 0$ のとき緩慢関数ではない。

補題 5.17 X が定理 5.16 (ii) を $X = X_1$ として満たすとする、次が成立する。

(1) $\forall \delta > 0$ に対して、定数 $C_\delta > 0$ と $x_\delta > 0$ が存在して、 $y \geq x_\delta$ に対して $P(|X| > y) \geq C_\delta y^{-(\alpha+\delta)}$ 。

特に、 $E[X^2] = \infty$ 。

(2) 定数 $C > 0$ と $x_0 > 0$ が存在して、 $x \geq x_0$ ならば $\int_0^x yP(|X| > y) dy \leq Cx^2P(|X| > x)$ 。

証明: (1) 仮定より、ある $x_\delta > 0$ が存在して、 $x \geq x_\delta$ ならば

$$\frac{P(|X| > 2x)}{P(|X| > x)} = 2^{-\alpha} \frac{L(2x)}{L(x)} \geq 2^{-(\alpha+\delta)}$$

とできる。よって、 $n \geq 1$ に対して、 $P(|X| > 2^n x_\delta) \geq P(|X| > x_\delta) 2^{-(\alpha+\delta)n}$ となる。ここで、 $\forall y \geq x_\delta$ に対して n を $2^{n-1} \leq y/x_\delta < 2^n$ ととれば、

$$P(|X| > y) \geq P(|X| > 2^n x_\delta) \geq P(|X| > x_\delta) 2^{-(\alpha+\delta)n} \geq C_\delta y^{-(\alpha+\delta)}$$

を得る。ただし $C_\delta = (x_\delta/2)^{\alpha+\delta}P(|X| > x_\delta)$ とした。次に、 $\delta > 0$ を $\alpha + \delta < 2$ と選べば、

$$E[X^2] = \int_0^\infty 2yP(|X| > y) dy \geq \int_{x_\delta}^\infty 2yC_\delta y^{-(\alpha+\delta)} dy = \left[\frac{2C_\delta}{2 - (\alpha + \delta)} y^{2-(\alpha+\delta)} \right]_{x_\delta}^\infty = \infty.$$

(2) $I(x) = \int_0^x yP(|X| > y) dy$ とおくと、 $t > s > 0$ に対して

$$I(tx) - I(sx) = \int_{sx}^{tx} yP(|X| > y) dy = x^2 \int_s^t uP(|X| > ux) du. \quad (5.26)$$

$\varepsilon > 0$ を任意とするとある $T > 0$ があって、 $u \geq T$ ならば $\frac{P(|X| > ux)}{P(|X| > u)} \geq (1 - \varepsilon)x^{-\alpha}$ とできるので、

$$I(tx) - I(sx) \geq (1 - \varepsilon)x^{2-\alpha} \int_s^t uP(|X| > u) du = (1 - \varepsilon)x^{2-\alpha}\{I(t) - I(s)\}, \quad t > s > T.$$

ここで、 $\lim_{t \rightarrow \infty} I(t) = \int_0^\infty yP(|X| > y) dy = E[|X|]/2 = \infty$ より、両辺を $I(t)$ で割って $t \rightarrow \infty$ とすると、

$$\liminf_{t \rightarrow \infty} \frac{I(tx)}{I(t)} \geq (1 - \varepsilon)x^{2-\alpha}, \quad \text{左辺は } \varepsilon > 0 \text{ によらないので} \quad \liminf_{t \rightarrow \infty} \frac{I(tx)}{I(t)} \geq x^{2-\alpha}. \quad (5.27)$$

(5.26) で $s = 1$ として両辺を $I(x)$ で割って変形すると、

$$\frac{I(tx)}{I(x)} = 1 + \frac{x^2 P(|X| > x)}{I(x)} \int_1^t u \frac{P(|X| > ux)}{P(|X| > x)} du \leq 1 + \frac{x^2 P(|X| > x)}{I(x)} \frac{t^2 - 1}{2}.$$

ここで、 $x \rightarrow \infty$ として、(5.27) より $t > 1$ のとき、

$$t^{2-\alpha} \leq \liminf_{x \rightarrow \infty} \frac{I(tx)}{I(x)} = 1 + \frac{t^2}{2} \liminf_{x \rightarrow \infty} \frac{x^2 P(|X| > x)}{I(x)}, \quad \text{すなわち、} \liminf_{x \rightarrow \infty} \frac{x^2 P(|X| > x)}{I(x)} \geq \frac{2(t^{2-\alpha} - 1)}{t^2}.$$

よって、 $x_0 > 0$ があって、 $x \geq x_0$ ならば $\frac{x^2 P(|X| > x)}{I(x)} \geq \frac{t^{2-\alpha} - 1}{t^2}$ となり、証明は完了した。 \square

定理 5.16 の証明: まず、 $\lim_{n \rightarrow \infty} nP(|X_1| > a_n) = 1$ に注意する。これは、定義より $nP(|X_1| > a_n) \leq 1$ であり、また、 $\forall \varepsilon > 0$ に対して条件 (ii) と a_n の取り方より

$$(1 + 2\varepsilon)^{-\alpha} = \lim_{n \rightarrow \infty} \frac{P(|X_1| > \frac{(1+2\varepsilon)a_n}{1+\varepsilon})}{P(|X_1| > \frac{a_n}{1+\varepsilon})} \leq \liminf_{n \rightarrow \infty} \frac{P(|X_1| > a_n)}{1/n}$$

となることから従う。よって、(ii) より、 $x > 0$ に対して、

$$nP(|X_1| > xa_n) = \frac{P(|X_1| > xa_n)}{P(|X_1| > a_n)} nP(|X_1| > a_n) \rightarrow x^{-\alpha}, \quad (5.28)$$

さらに、(i) と組み合わせると、 $x > 0$ に対して、

$$nP(X_1 > xa_n) = \frac{P(X_1 > xa_n)}{P(|X_1| > xa_n)} nP(|X_1| > xa_n) \rightarrow \theta x^{-\alpha}, \quad (5.29)$$

同様に、 $x < 0$ に対して $nP(X_1 < xa_n) \rightarrow (1 - \theta)|x|^{-\alpha}$ を得る。

$\varepsilon > 0$ を任意とし、 $I_n(\varepsilon) = \{k \in \mathbf{N}; 1 \leq k \leq n, |X_k| > \varepsilon a_n\}$ とし、更に、

$$\begin{aligned} \mu_1(\varepsilon) &= E[X_1 1_{\{\varepsilon a_n < |X_1| \leq a_n\}}], & \mu_2(\varepsilon) &= E[X_1 1_{\{|X_1| \leq \varepsilon a_n\}}], \\ S_n(\varepsilon) &= \sum_{k \in I_n(\varepsilon)} X_k, & T_n(\varepsilon) &= S_n - b_n - (S_n(\varepsilon) - n\mu_1(\varepsilon)) = \sum_{k=1}^n \{X_k 1_{\{|X_k| \leq \varepsilon a_n\}} - \mu_2(\varepsilon)\}, \end{aligned}$$

とおく。

$T_n(\varepsilon)$ について、 $X_k^\varepsilon = X_k 1_{\{|X_k| \leq \varepsilon a_n\}}$ とおくと、

$$\begin{aligned} E[(T_n(\varepsilon)/a_n)^2] &= V(T_n(\varepsilon)/a_n) = nV(X_1^\varepsilon/a_n) \leq nE[(X_1^\varepsilon/a_n)^2] = n \int_0^\infty 2xP(|X_1^\varepsilon/a_n| > x) dx \\ &\leq 2n \int_0^\varepsilon xP(|X_1| > xa_n) dx = \frac{2n}{a_n^2} \int_0^{\varepsilon a_n} yP(|X_1| > y) dy \end{aligned}$$

であるが、 $\varepsilon a_n \rightarrow \infty$ に注意して補題 5.17 (2) と (5.28) を用いて、次を得る。

$$E[(T_n(\varepsilon)/a_n)^2] \leq \frac{2n}{a_n^2} C(\varepsilon a_n)^2 P(|X_1| > \varepsilon a_n) \rightarrow 2C\varepsilon^{2-\alpha}. \quad (5.30)$$

$S_n(\varepsilon)$ について、 $\#I_n(\varepsilon) = k$ の条件のもと、 $S_n(\varepsilon)/a_n$ の分布は $F_n^\varepsilon(x) = P(X_1/a_n \leq x | |X_1|/a_n > \varepsilon)$ を共通の分布関数とする k 個の独立な確率変数の和として表されることに注意する。また、 $\#I_n(\varepsilon)$ は二項分布 $B(n, p_n^\varepsilon)$, $p_n^\varepsilon = P(|X_1| > \varepsilon a_n)$, に従うので、 F_n^ε の特性関数を $\psi_n^\varepsilon(t)$ とすると、

$$E[e^{itS_n(\varepsilon)/a_n}] = E\left[E[e^{itS_n(\varepsilon)/a_n} | \#I_n^\varepsilon]\right] = \sum_{k=0}^n (\psi_n^\varepsilon(t))^k \binom{n}{k} (p_n^\varepsilon)^k (1-p_n^\varepsilon)^{n-k} = \{1 + (\psi_n^\varepsilon(t) - 1)p_n^\varepsilon\}^n.$$

ここで、

$$F_n^\varepsilon(x) = \begin{cases} 1 - \frac{P(X_1 > xa_n)}{P(|X_1| > \varepsilon a_n)} & \rightarrow 1 - \theta(x/\varepsilon)^{-\alpha}, & x > \varepsilon, \\ \frac{P(X_1 \leq xa_n)}{P(|X_1| > \varepsilon a_n)} & \rightarrow \theta(|x|/\varepsilon)^{-\alpha}, & x < -\varepsilon, \end{cases}$$

より、

$$\psi_n^\varepsilon(t) \rightarrow \psi^\varepsilon(t) := \varepsilon^\alpha \int_\varepsilon^\infty e^{itx} \theta \alpha \frac{dx}{x^{\alpha+1}} + \varepsilon^\alpha \int_{-\infty}^{-\varepsilon} e^{itx} (1-\theta) \alpha \frac{dx}{|x|^{\alpha+1}}$$

となることと、 $np_n^\varepsilon \rightarrow \varepsilon^{-\alpha}$ および $\int_\varepsilon^\infty \alpha x^{-\alpha-1} dx = \varepsilon^{-\alpha}$ に注意して、

$$E[e^{itS_n(\varepsilon)/a_n}] \rightarrow e^{\varepsilon^{-\alpha}(\psi^\varepsilon(t)-1)} = \exp\left(\int_\varepsilon^\infty (e^{itx} - 1) \theta \alpha \frac{dx}{x^{\alpha+1}} + \int_{-\infty}^{-\varepsilon} (e^{itx} - 1) (1-\theta) \alpha \frac{dx}{|x|^{\alpha+1}}\right)$$

を示すことができる。また、(5.29) より、

$$n\mu_1(\varepsilon)/a_n = nE[(X_1/a_n) 1_{\{\varepsilon a_n < |X_1| \leq a_n\}}] \rightarrow \int_\varepsilon^1 x \theta \alpha \frac{dx}{x^{\alpha+1}} + \int_{-1}^{-\varepsilon} x(1-\theta) \alpha \frac{dx}{|x|^{\alpha+1}}$$

を得る。従って、 $E[\exp\{it(S_n(\varepsilon) - n\mu_1(\varepsilon))/a_n\}] \rightarrow$

$$\begin{aligned} &\exp\left\{\int_1^\infty (e^{itx} - 1) \theta \alpha \frac{dx}{x^{\alpha+1}} + \int_\varepsilon^1 (e^{itx} - 1 - itx) \theta \alpha \frac{dx}{x^{\alpha+1}} \right. \\ &\quad \left. + \int_{-\infty}^{-1} (e^{itx} - 1) (1-\theta) \alpha \frac{dx}{|x|^{\alpha+1}} + \int_{-1}^{-\varepsilon} (e^{itx} - 1 - itx) (1-\theta) \alpha \frac{dx}{|x|^{\alpha+1}}\right\} \end{aligned} \quad (5.31)$$

となる。ここで、 $e^{itx} - 1 - itx \sim -t^2 x^2/2$ ($x \rightarrow 0$) より、 $0 < \alpha < 2$ のとき、

$$\int_0^1 (e^{itx} - 1 - itx) \frac{dx}{x^{\alpha+1}}, \quad \int_{-1}^0 (e^{itx} - 1 - itx) \frac{dx}{|x|^{\alpha+1}}$$

はともに収束する。よって、 $\varepsilon \rightarrow 0$ として (5.31) は

$$\exp\left\{i\beta t + \int_0^\infty \left(e^{itx} - 1 - \frac{itx}{1+x^2}\right) \theta \alpha \frac{dx}{x^{1+\alpha}} + \int_{-\infty}^0 \left(e^{itx} - 1 - \frac{itx}{1+x^2}\right) (1-\theta) \alpha \frac{dx}{|x|^{1+\alpha}}\right\}, \quad (5.32)$$

$$\text{ただし、 } \beta = (2\theta - 1)\alpha \left\{ \int_1^\infty \frac{x}{1+x^2} \frac{dx}{x^{1+\alpha}} + \int_0^1 \left(\frac{x}{1+x^2} - x\right) \frac{dx}{x^{1+\alpha}} \right\} \quad (5.33)$$

に収束する。

定理 5.16 の証明を完了させるために次の補題を準備する。

補題 5.18 各 $\varepsilon > 0$ で $h_n(\varepsilon) \rightarrow h(\varepsilon)$, $g_n(\varepsilon) \rightarrow g(\varepsilon)$ ($n \rightarrow \infty$) かつ $h(\varepsilon) \rightarrow h(0)$, $g(\varepsilon) \rightarrow g(0)$ ($\varepsilon \rightarrow 0$) であれば、 $\{\varepsilon_n\}$ を $h_n(\varepsilon_n) \rightarrow h(0)$ かつ $g_n(\varepsilon_n) \rightarrow g(0)$ となるように選ぶことができる。

証明: $N_1 < N_2 < \dots$ を、 $n \geq N_m$ ならば $|h_n(1/m) - h(1/m)| \leq 1/m$ かつ $|g_n(1/m) - g(1/m)| \leq 1/m$ となるように選べる。ここで、 $n < N_1$ のとき $\varepsilon_n = 1$, $N_m \leq n < N_{m+1}$ のときは $\varepsilon_n = 1/m$ と選べば、

$$|h_n(\varepsilon_n) - h(0)| \leq |h_n(1/m) - h(1/m)| + |h(1/m) - h(0)| \leq 1/m + |h(1/m) - h(0)|$$

で、 $n \rightarrow \infty$ のとき $m \rightarrow \infty$ となるので、 $h_n(\varepsilon_n) \rightarrow h(0)$ となる。 $g_n(\varepsilon_n) \rightarrow g(0)$ も同様に示せる。 \square

定理 5.16 の証明の続き: 補題 5.18 を $h_n(\varepsilon) = E[(T_n(\varepsilon)/a_n)^2]$, $g_n(\varepsilon) = E[\exp\{it(S_n(\varepsilon) - n\mu_1(\varepsilon))/a_n\}]$ として用いると、 $\{\varepsilon_n\}$ が存在して、(5.30) より $h_n(\varepsilon_n) \rightarrow 0$, また $g_n(\varepsilon) \rightarrow (5.32)$ とできる。これより、 $T_n(\varepsilon_n)/a_n \rightarrow 0$ in L^2 , 特に、 $T_n(\varepsilon_n)/a_n \rightarrow 0$ in prob. であり、一方、(5.32) は $t = 0$ で連続なので確率統計学 II 定理 6.30 (2) より (5.32) はある確率変数 Y の特性関数であり、 $(S_n(\varepsilon_n) - n\mu_1(\varepsilon_n))/a_n \rightarrow Y$ in law となることから従う。以上より、

$$\frac{S_n - b_n}{a_n} = \frac{T_n(\varepsilon_n)}{a_n} + \frac{S_n(\varepsilon_n) - n\mu_1(\varepsilon_n)}{a_n} \rightarrow Y \quad \text{in law}$$

を得る。(一般に、 $X_n \rightarrow 0$ in prob. かつ $Y_n \rightarrow Y$ in law であれば、 $X_n + Y_n \rightarrow Y$ in law となることを用いた。) \square

定義 5.5 (吸引域) X_k の分布関数を F とする。ある定数 $a_n > 0$, b_n が存在して、 $\frac{S_n - b_n}{a_n}$ が指数 $\alpha \in (0, 2)$ の安定分布 Z に法則収束するとき、 $F(x)$ あるいは X_k の分布は指数 α の安定分布 Z の**吸引域**に属するという。

定理 5.16 は X_i の分布が (i), (ii) を満たせば指数 α の吸引域に属することを意味している。

安定分布の理論はより広い族である無限分解可能分布や対応する確率過程である Lévy 過程 (独立増分を持ち時間的一様 $X_{t+s} - X_t \sim X_s$ となる、例えば Brown 運動や Poisson 過程、複合 Poisson 過程など) と深く関連しており、詳しく調べられている。その教科書としては例えば次がある。

[S] 佐藤 健一: 加法過程, 紀伊國屋書店, 1999. (より新しい内容を含む著者による英語版もある。)