

第1回 確率過程と時系列 (2.1, 2.3.1, A.1.1)

村澤 康友

2020年9月29日

今日のポイント

1. 試行の結果によって値が決まる数列を確率変数列 (離散確率過程) という。確率変数列の実現値を時系列という。
2. 必要なら時系列を対数系列, 差分 (階差) 系列, 対数差分 (階差) 系列に変換する。対数差分は変化率と近似的に等しい。
3. 時系列 (y_1, \dots, y_T) から推定した回帰式 $E(Y_t|Y_{t-1})$ は, 一定の条件の下で Y_{T+1} の予測に使える。

目次

1	確率過程	1
1.1	確率	1
1.2	確率変数 (p. 188)	1
1.3	確率ベクトル	1
1.4	確率変数列	1
2	時系列	2
2.1	時系列 (p. 36)	2
2.2	時系列の変換 (p. 37)	2
3	時系列予測	2
3.1	同時分布と周辺分布	2
3.2	条件つき分布	4
3.3	回帰 (p. 25)	4
3.4	1期先予測	4
4	今日のキーワード	5
5	次回までの準備	5

1 確率過程

1.1 確率

定義 1. 結果が偶然に支配される実験を試行という。

定義 2. 試行において起こりうる結果を標本点という。

定義 3. 標本点全体の集合を標本空間という。

注 1. 標本点を ω , 標本空間を Ω で表すことが多い。

定義 4. 標本空間の部分集合を事象という。

定義 5. 事象に対して定義され, 以下の公理を満たす関数 $P(\cdot)$ を確率という。

$$1. 0 \leq P(\cdot) \leq 1$$

$$2. P(\Omega) = 1$$

3. (σ 加法性) A_1, A_2, \dots が排反なら

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

1.2 確率変数 (p. 188)

定義 6. 試行の結果によって値が決まる変数を確率変数という。

注 2. すなわち $X: \Omega \rightarrow \mathbb{R}$.

1.3 確率ベクトル

定義 7. 試行の結果によって値が決まるベクトルを確率ベクトルという。

注 3. すなわち $(X_1, \dots, X_n): \Omega \rightarrow \mathbb{R}^n$.

1.4 確率変数列

定義 8. 試行の結果によって値が決まる数列を確率変数列 (離散確率過程) という。

注 4. すなわち $\{X_t\} : \Omega \rightarrow \mathbb{R}^\infty$. 確率変数・確率ベクトルと同様に, 起こりうる結果に対して確率を定義できる.

2 時系列

2.1 時系列 (p. 36)

確率変数列 $\{X_t\}$ の実現値 $\{x_t\}$ のうち, 実際に観測される部分を (x_1, \dots, x_T) とする.

定義 9. 確率変数列の実現値を時系列という.

定義 10. 時系列の観測値の数を時系列の長さという.

注 5. 時系列の長さ \neq 標本の大きさ. 1 度しか観測されない時系列の標本の大きさは 1. したがって (x_1, \dots, x_T) は長さ T の時系列の大きさ 1 の標本.

2.2 時系列の変換 (p. 37)

必要なら分析の前に時系列を変換する.

定義 11. $\{x_t\}$ の対数系列は $\{\ln x_t\}$.

注 6. 自然対数で変換する.

注 7. $x_t > 0$ でないと変換できない.

例 1. 株価指数 (NYSE 総合指数) の原系列と対数系列 (図 1).

定義 12. $\{x_t\}$ の差分 (階差) 系列は $\{\Delta x_t\}$.

注 8. Δ は (後退) 差分演算子. すなわち $\Delta x_t := x_t - x_{t-1}$.

定義 13. $\{x_t\}$ の変化 (成長) 率系列は $\{\Delta x_t / x_{t-1}\}$.

注 9. 時系列分析ではあまり使われない.

定義 14. $\{x_t\}$ の対数差分 (階差) 系列は $\{\Delta \ln x_t\}$.

注 10. 時系列分析では変化率の代わりによく使われる.

例 2. 株価指数 (NYSE 総合指数) の差分系列と対数差分系列 (図 2).

補題 1. $x = 0$ の近傍において

$$\ln(1+x) \approx x$$

証明. マクローリン展開より

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

したがって $x = 0$ の近傍において

$$e^x \approx 1 + x$$

両辺を対数変換すると

$$x \approx \ln(1+x)$$

□

定理 1.

$$\Delta \ln x_t \approx \frac{\Delta x_t}{x_{t-1}}$$

証明. 補題より

$$\begin{aligned} \Delta \ln x_t &:= \ln x_t - \ln x_{t-1} \\ &= \ln \frac{x_t}{x_{t-1}} \\ &= \ln \left(1 + \frac{x_t - x_{t-1}}{x_{t-1}} \right) \\ &\approx \frac{x_t - x_{t-1}}{x_{t-1}} \end{aligned}$$

□

注 11. 対数差分は正と負が対称. 変化率は正と負が非対称.

例 3. $x_1 := 1$ すなわち $\ln x_1 = 0$ とする.

1. $\Delta \ln x_2 := 1$, $\Delta \ln x_3 := -1$ なら $\ln x_3 = 0$ より $x_3 = 1 = x_1$.
2. $\Delta x_2 / x_1 = 1$, $\Delta x_3 / x_2 = -1$ なら 100% 増の 100% 減だから $x_3 = 0 \neq x_1$.

3 時系列予測

3.1 同時分布と周辺分布

(X, Y) を確率ベクトルとする.

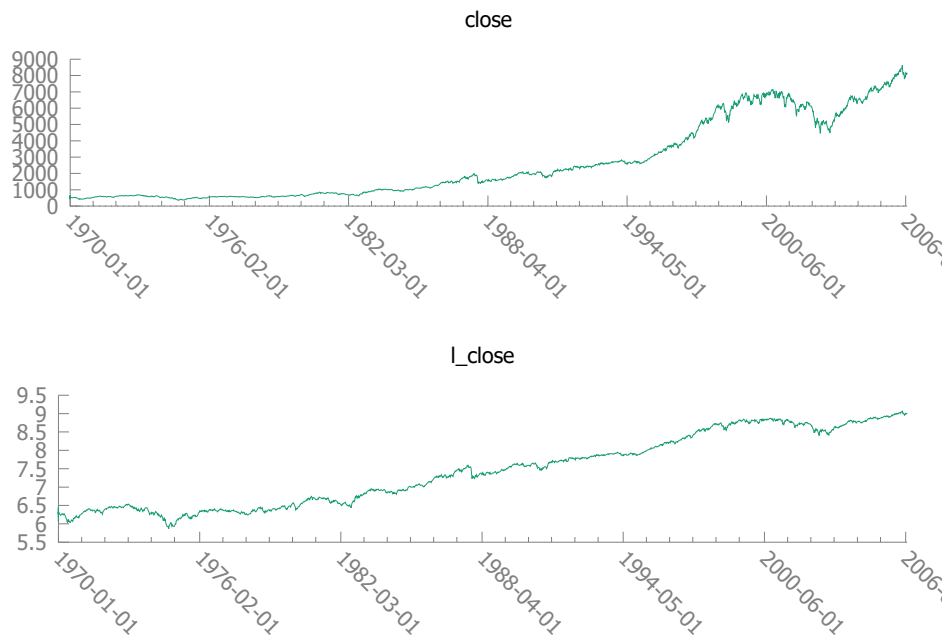


図1 NYSE 総合指数（週次）の原系列と対数系列

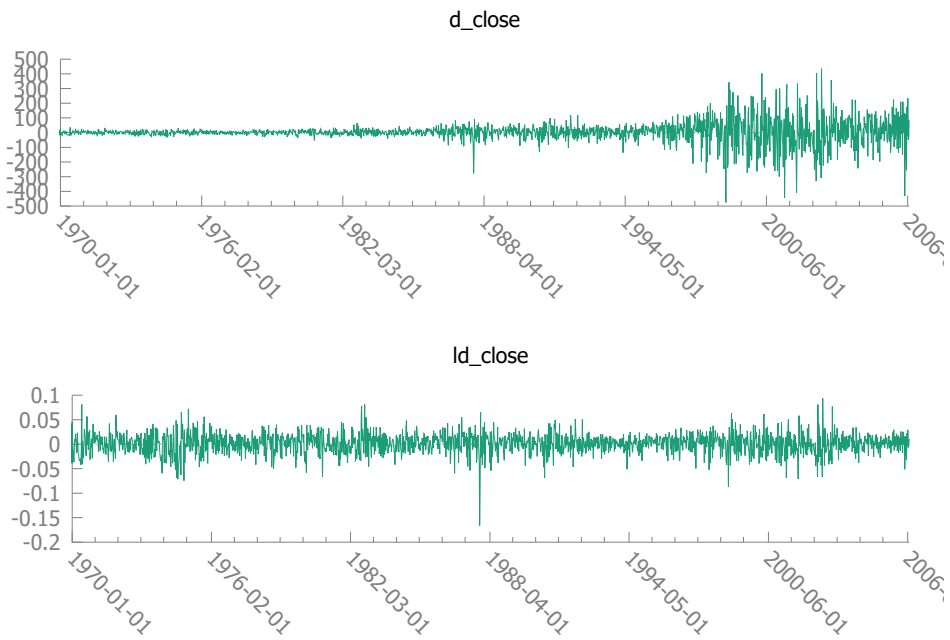


図2 NYSE 総合指数（週次）の差分系列と対数差分系列

定義 15. (X, Y) の同時 (結合) 累積分布関数 (*cumulative distribution function, cdf*) は, 任意の (x, y) について

$$F_{X,Y}(x, y) := \Pr[X \leq x, Y \leq y]$$

定義 16. X の周辺 *cdf* は, 任意の x について

$$F_X(x) := \Pr[X \leq x]$$

定義 17. (X, Y) の同時 (結合) 確率質量関数 (*probability mass function, pmf*) は, 任意の (x, y) について

$$p_{X,Y}(x, y) := \Pr[X = x, Y = y]$$

定義 18. X の周辺 *pmf* は, 任意の x について

$$p_X(x) := \Pr[X = x]$$

注 12. 同時 *pmf* と周辺 *pmf* の関係は

$$p_X(x) = \sum_y p_{X,Y}(x, y)$$

定義 19. 任意の (x, y) について

$$F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(s, t) ds dt$$

となる $f_{X,Y}(\cdot, \cdot)$ を (X, Y) の同時 (結合) 確率密度関数 (*probability density function, pdf*) という.

注 13. $F_{X,Y}(\cdot, \cdot)$ が微分可能なら

$$f_{X,Y}(x, y) = \frac{\partial^2 F_{X,Y}}{\partial x \partial y}(x, y)$$

定義 20. X の周辺 *pdf* は, 任意の x について

$$f_X(x) := \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy$$

3.2 条件つき分布

定義 21. $X = x$ が与えられたときの Y の条件つき *pmf* は, 任意の y について

$$p_{Y|X}(y|X = x) := \frac{p_{X,Y}(x, y)}{p_X(x)}$$

注 14. 条件つき確率で定義する.

定義 22. $X = x$ が与えられたときの Y の条件つき *pdf* は, 任意の y について

$$f_{Y|X}(y|X = x) := \frac{f_{X,Y}(x, y)}{f_X(x)}$$

注 15. 条件つき確率と同様に定義する.

定義 23. $X = x$ が与えられたときの Y の条件つき期待値は

$$E(Y|X = x) := \begin{cases} \sum_y y p_{Y|X}(y|X = x) & (\text{離散}) \\ \int_{-\infty}^{\infty} y f_{Y|X}(y|X = x) dy & (\text{連続}) \end{cases}$$

3.3 回帰 (p. 25)

定義 24. $E(Y|X)$ を求めることを, Y を X に回帰するという.

定義 25. $E(Y|X)$ を与える式を, Y の X 上への回帰モデル (回帰式, 回帰関数) という.

注 16. すなわち

$$E(Y|X) = r(X)$$

定義 26. 線形な回帰モデルを線形回帰モデルという.

注 17. すなわち

$$E(Y|X) = \alpha + \beta X$$

3.4 1期先予測

確率変数列 $\{Y_t\}$ の実現値 (y_1, \dots, y_T) から Y_{T+1} を予測したい. Y_{T+1} の (Y_1, \dots, Y_T) 上への線形回帰モデルは

$$E(Y_{T+1}|Y_1, \dots, Y_T) = \alpha + \beta_1 Y_1 + \dots + \beta_T Y_T$$

Y_{T+1} が観測されていないため, この式は推定できない.

$t = 2, \dots, T$ について, Y_t の Y_{t-1} 上への単回帰モデルを仮定する. すなわち

$$E(Y_t|Y_{t-1}) = \alpha + \beta Y_{t-1}$$

この式は観測値 (y_1, \dots, y_T) から推定できる. 一定の条件の下で, この式を Y_{T+1} の予測に使うことができる.

4 今日のキーワード

試行, 標本点, 標本空間, 事象, 確率, 確率変数, 確率ベクトル, 確率変数列 (離散確率過程), 時系列, (時系列の) 長さ, 対数系列, 差分 (階差) 系列, 変化 (成長) 率系列, 対数差分 (階差) 系列, 同時 (結合) 累積分布関数 (cdf), 周辺 cdf, 同時 (結合) 確率質量関数 (pmf), 周辺 pmf, 同時 (結合) 確率密度関数 (pdf), 周辺 pdf, 条件つき pmf, 条件つき pdf, 条件つき期待値, 回帰, 回帰モデル (回帰式, 回帰関数), 線形回帰モデル

5 次回までの準備

提出 宿題 1

復習 教科書第 2 章 1 節, 第 2 章 3.1 節, 補論 A.1.1 節, 復習テスト 1

予習 教科書第 2 章 3.1 節, 第 4 章 3.3 節