

尤度・最尤推定・単回帰分析

樋口さぶろお <https://hig3.net>

龍谷大学 先端理工学部 数理・情報科学課程

多変量解析☆演習 L00(2026-09-24 Thu)

最終更新: Time-stamp: "2026-09-20 Sun 07:35 JST hig"

今日の目標

- L01 科目ののりを説明できる, Moodle 使える
- 尤度を説明できる. 最尤推定できる
- 単回帰の確率モデルと最尤推定を説明できる



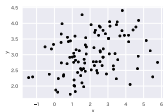
ここまで来たよ

- はじめに
 - この授業どんなのり?

科目の目標

時系列データ, 2次元以上の実数値データと, 問題が与えられたときに, 問題を数学・統計学的に定式化し, 統計手法を選択し, 統計ソフトウェアを使って データを分析し, 結論を述べることができる.

- 2次元以上の実数値データ: **多変量解析** = 確率統計Ⅰの, 2次元分布の続きのようなもの. 入力 → 出力. データからモデルのパラメタや新しい入力に対する出力を推定する (推測統計).



- 時系列データ: **時系列解析** = 確率過程 (時間とともに変化する確率変数). データからモデルのパラメタや未来の値を推定する (推測統計).



- 統計ソフトウェア: Google Colaboratory=Python + Jupyter Notebook

学習目標

講義概要 → シラバス

現実世界の現象を理解し、数理モデルとの関係を明らかにするためには、観察・実験により取得したデータを整理・解析することが必要である。限られたデータから数理モデルのパラメタを推測する推測統計と、必要な確率論を説明する。

到達目標 → シラバス

構造的 いくつかの典型的な多次元データ分析，時系列データの手法について，その問題設定を含めて説明できる．

実習的 多次元データ，時系列データが与えられたときに，適切な手法を選択して分析を実行し，結論を述べられる．

関係する科目

プログラム「データサイエンス」学修プログラム．（数理）データサイエンス履修モデル．

前

- データ分析 Web
- **確率統計 I** Gemini Notebook，参考文献：公式と例題で学ぶ統計学入門
- 線形代数☆演習 I,II 改訂新版すぐわかる線形代数

重なるの多い科目

- （知能）データインテリジェンス

後

- 機械学習 I,II(3 年前期)

尤度・最尤推定・単回帰分析

樋口さぶろお <https://hig3.net>

龍谷大学 先端理工学部 数理・情報科学課程

多変量解析☆演習 L01(2026-09-24 Thu)

最終更新: Time-stamp: "2026-09-20 Sun 07:35 JST hig"

今日の目標

- (従来ののりで) 回帰分析を説明できる
- 正規分布を説明できる
- 尤度を説明できる．最尤推定できる



ここまで来たよ

1. 尤度・最尤推定・単回帰分析

- 回帰分析の復習
- 正規分布の復習
- 尤度
- 最尤推定

回帰分析（単回帰分析）の復習

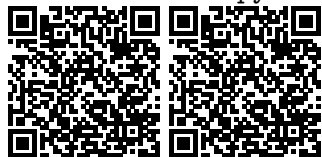
データ (x_i, y_i) ($i = 1, \dots, N$) から、回帰係数を、最小 2 乗法で決めたのだった.

[データ分析 \(2025\)ex06](#)[データ分析 \(2025\)ex07](#)[データ分析 \(2025\)ex08](#)

Notebooks

https://github.com/takatakamanbou/Data/blob/2025/Data2025_ex07notebookA.ipynb

https://github.com/takatakamanbou/Data/blob/2025/Data2025_ex07notebookA.ipynb



$$y = \hat{\beta}_1 x + \hat{\beta}_0.$$

データ分析 (2025)ex の記号では, $a \leftrightarrow \hat{\beta}_1$, $b \leftrightarrow \hat{\beta}_0$.

$$\hat{\beta}_1 = \frac{s_{xy}}{s_{xx}},$$

$$\hat{\beta}_0 = \bar{y} - \frac{s_{xy}}{s_{xx}} \bar{x}$$

$$y = \frac{s_{xy}}{s_{xx}} x + \bar{y} - \frac{s_{xy}}{s_{xx}} \bar{x}$$

$$y - \bar{y} = \frac{s_{xy}}{s_{xx}} (x - \bar{x})$$

ここで,

$$\bar{x} = \frac{1}{n} \sum_i x_i \quad \text{平均値っぽい形}$$

$$\bar{y} = \frac{1}{n} \sum_i y_i \quad \text{平均値っぽい形}$$

$$s_{xy} = \frac{1}{n} \sum_i x_i y_i - \bar{x} \cdot \bar{y} = \frac{1}{n} \sum_i (x_i - \bar{x})(y_i - \bar{y}) \quad \text{共分散っぽい形}$$

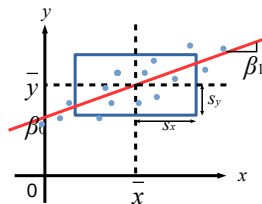
$$s_{xx} = \frac{1}{n} \sum_i x_i^2 - \bar{x}^2 = \frac{1}{n} \sum_i (x_i - \bar{x})^2 \quad \text{分散っぽい形}$$

$y = \hat{\beta}_0 + \hat{\beta}_1 \cdot x$ ($\hat{\beta}_0, \hat{\beta}_1$: 回帰係数) に代入.

定理

回帰直線 久保川 統計学入門 §12 推定結果 $\hat{\beta}_0, \hat{\beta}_1$ を係数とする xy 平面の直線

$$y - \bar{y} = \frac{s_{xy}}{s_{xx}}(x - \bar{x})$$



回帰係数, 予測値の信頼区間

確率統計 II,III

ここまで来たよ

1. 尤度・最尤推定・単回帰分析

- 回帰分析の復習
- 正規分布の復習
- 尤度
- 最尤推定

標準正規分布 standard normal distribution $N(0, 1^2)$ の性質

定義 (標準正規分布 $N(0, 1^2)$) 久保川 統計学入門 式 (5.2)

次の確率密度関数を持つ確率変数 Z を，標準正規分布 $N(0, 1^2)$ にしたかうという．

$$f_Z(z) = \phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}.$$

累積分布関数

$$F_Z(z) = \Phi(z) = \int_{-\infty}^z \phi(z') dz'.$$

ϕ, Φ は f, F のギリシャ文字．あまりに有名分布なので専用文字を使う． $\Phi(z)$ の積分は具体的に書けない．

Python で `scipy.stats.norm(loc=0,scale=1)`

Excel で `=norm.s.dist(z,FALSE)`

一般の正規分布 $N(b, a^2)$

$Z \sim N(0, 1^2)$ に対して, $X = aZ + b$ を考える.

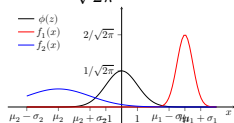
久保川 統計学入門 公式 5.17

$$E[X^0] = 1,$$

$$\mu = E[X] = E[aZ + b] = b,$$

$$\sigma^2 = \text{Var}(X) = \text{Var}(aZ + b) = a^2.$$

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \rightsquigarrow f_X(x) = \frac{1}{a} f_Z\left(\frac{x-b}{a}\right) = \frac{1}{\sqrt{2\pi}a} e^{-\frac{(x-b)^2}{2a^2}}$$



定義 (一般の正規分布 $N(\mu, \sigma^2)$)

久保川 統計学入門 式 (5.1)

次の確率密度関数を持つ確率変数 X を, 母平均値 μ , 母分散 σ^2 の正規分布 $N(\mu, \sigma^2)$ (normal distribution) にしたがるという.

$$f_X(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Python で `scipy.stats.norm(loc=b, scale=a)`.

$X \sim N(b, a^2)$ を標準化すると $Z \sim N(0, 1^2)$.

ここまで来たよ

1. 尤度・最尤推定・単回帰分析

- 回帰分析の復習
- 正規分布の復習
- 尤度
- 最尤推定

正規モデル（母平均値パラメタ）

ドーナツ製造機は、母平均値 μ g, 母分散 σ^2 g² の正規分布にしたがう質量 x g のドーナツを作り出す.

μ, σ^2 のような、母分布を決める数を母数, パラメタという（いまたまた母平均値と母分散）よく θ と書く.

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

ドーナツの標本 $x_1, \dots, x_N$ が得られる確率は、標本抽出は独立だから、

$$f_N(x_1, \dots, x_N|\mu, \sigma^2) = \prod_{i=1}^N f(x_i|\mu, \sigma^2) = (2\pi\sigma^2)^{-N/2} e^{-\sum \frac{(x_i-\mu)^2}{2\sigma^2}}.$$

$x_1, \dots, x_N$ を変数, μ, σ^2 を定数と思うと, $f(x_1, \dots, x_N)$ は N 変数の確率密度関数.

もうドーナツは焼けたのだから, $x_1, \dots, x_N$ を定数, μ, σ^2 を未知の定数 (母数) と思うと, $L(\mu, \sigma^2) = f(x_1, \dots, x_N|\mu, \sigma^2)$ を尤度という.

大注意; 尤度の値は確率だが, 尤度は確率関数, 確率密度関数ではない.

母数について積分して1になるわけではない

ここまで来たよ

1. 尤度・最尤推定・単回帰分析

- 回帰分析の復習
- 正規分布の復習
- 尤度
- 最尤推定

標本から母数を推定する

母平均値, 母分散のような母数には, それぞれ専用の推定量の「公式」があった.

確率統計 I(2026)L10

- 母平均値 ← 標本母平均値
- 母分散 ← **不偏**標本分散
- $E[X^4]$?

でも, 母数は確率密度関数に入ってるパラメタならなんでもいいんだから, 公式集にきりなくない?

どんな母数も推定できる最尤推定

定義 (最尤推定 (MLE, Maximum Likelihood Estimation))

標本に対して, 母数を, 尤度が最大になる値として推定すること

最大 $\leadsto$ 停留点

$$\frac{\partial L}{\partial \mu} = \frac{\partial L}{\partial \sigma^2} = 0.$$

不偏標本分散 $\neq$ 母分散の最尤推定値

L01-Q1

TA Prob and Sol: 正規分布の母数の最尤推定

未知の母平均値 μ , 母分散 σ^2 の正規分布

$$f(x|\theta) = f(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

からサイズ N の標本 $\{x_1, \dots, x_N\}$ を得た.

対数尤度は $\log L(\theta) = \sum_{i=1}^N \log f(x_i|\mu, \sigma)$ である.

- ① $N = 2$ のとき, 対数尤度を最大化することにより μ, σ^2 を最尤推定しよう.
- ② 一般の N に対して, 対数尤度を最大化することにより μ, σ^2 を最尤推定しよう.

略解

$\log L(\mu, \sigma^2) = \sum_i -\frac{1}{2}(2\pi\sigma^2) - \frac{1}{2\sigma^2}(x_i - \mu)^2$ に対して,
 $\frac{\partial}{\partial \mu} \log L = \frac{\partial}{\partial (\sigma^2)} \log L = 0$ を考える.

$$\textcircled{1} \quad \mu = \frac{1}{2}(x_1 + x_2), \quad \sigma^2 = \left(\frac{x_1 - x_2}{2} \right)^2.$$

$$\textcircled{2} \quad \mu = \frac{1}{N}(x_1 + \cdots + x_N), \quad \sigma^2 = \frac{1}{N} \sum_{i=1}^N \left(x_i - \frac{1}{N}(x_1 + \cdots + x_N) \right)^2.$$

($1/(N-1)$ ではない).

次回予告：回帰係数も最尤推定で決められる