母数の最尤推定・ベイズ推定

樋口さぶろお https://hig3.net

龍谷大学大学院 先端理工学研究科 数理・情報科学コース

計算科学特論 L02(2026-04-20 Mon)

最終更新: Time-stamp: "2026-04-21 Tue 11:42 JST hig"

今日の目標

- Colab, PyMC で確率分布を可視化できる
- 最尤推定 (MLE) ができる
- 母数のベイズ推定 (最大事後確率 MAP) ができる

L01-Q1

Quiz 解答: 離散-連続混合型分布の周辺分布

1. $X \sim \mathrm{U}(0,1)$.
2. $Y \sim \mathrm{B}(1, \frac{1}{2})$.
3. $f_{\mathrm{X}|\mathrm{Y}}(x|1) = 2x \cdot u(x; 0, 1)$.
4. $Y \sim \mathrm{B}(1, \frac{7}{10})$.
5. まず X をルーレット $u(x; 0, 1)$ で得て，次にコイン $\mathrm{B}(1, x)$ で得るのが簡単.

PyMC

PyMC は，$f(y|x_1, \ldots, x_N)$ (複雑だが式は書ける) の値を計算したりグラフを描いたりはしないが，多数の y をサンプルすることができる．それでヒストグラムを描くことができる．そこから，代表値や代表的な部分集合を求めることができる．
サンプルする方法は MCMC(Markov Chain Monte Carlo 法).
ここでは，Colab で例を示す．

ここまで来たよ

1. 母比率の推定再訪・離散-連続型確率変数
2. 母数の最尤推定・ベイズ推定
   - 従来の母比率の点推定
   - 尤度と最尤推定
   - ベイズ推定

頻度主義での母比率の区間推定

ベルヌイ分布 $X \sim \mathrm{B}(1, p)$.

$$p(x) = \begin{cases} 1-p & (x=0) \\ p & (x=1) \end{cases}$$

パラメタ p は母平均値に等しい $\mathrm{E}[X] = p$.
標本 $X_1, \dots, X_N \sim \mathrm{B}(1, p)$ (独立同分布) からパラメタ p を推定したい

確率統計 I(2023)L13 の方針

母平均値の推定量として標本平均値 $\overline{X} = \frac{1}{N}(X_1 + \cdots + X_N)$ がある（一般論）のでこれを使おう.
標本平均値は母平均値に対する不偏推定量 $\mathrm{E}[\overline{X}] = \mathrm{E}[X]$.
疑問 母平均値でない母数の推定は？ 例: 母分散，$\mathrm{U}(c, d)$ の c, $\mathrm{B}(1, p)$ の p^2 や 母分散 $p(1-p)$. べつべつの理論が必要？

ここまで来たよ

1. 母比率の推定再訪・離散-連続型確率変数
2. 母数の最尤推定・ベイズ推定
   - 従来の母比率の点推定
   - 尤度と最尤推定
   - ベイズ推定

尤度 赤石 p.55

母数 p のベルヌイ分布 $X \sim \mathrm{B}(1,p)$ で，実現値 $x=0,1$ が得られる確率は，

$$P(X=x)=p^x(1-p)^{1-x}$$

サイズ N の標本の実現値 $x_1,\dots,x_N$ が得られる確率は，$X_1,\dots,X_N$ は独立同分布にしたがうので，$\sum_{i=1}^N x_i = k$ とおくと，

$$\begin{aligned}
P(X_1=x_1,X_2=x_2,\dots) =& P(X_1=x_1)\cdot P(X_2=x_2)\dots \\
=& p^{x_1}(1-p)^{1-x_1}\cdot p^{x_2}(1-p)^{1-x_2}\dots \\
=& p^{\sum_{i=1}^N x_i}(1-p)^{N-\sum_{i=1}^N x_i} \\
=& p^k(1-p)^{N-k}
\end{aligned}$$

これを，母数の関数 $L(p)$ とみなし，標本 $x_1,\dots,x_N$ が得られている（これらは定数とみなす）ときの尤度 (ゆうど)(likelyhood) という．$\log L(p)$ を対数尤度 赤石 p.58 という．

尤度=母数が「尤もらしい（もっともらしい）」度

尤度は，何かの確率変数が実現値 p をとる確率ではない

$\int L(p)\,\mathrm{d}p = 1$ でもない

最尤推定

母数 θ (あるいは，$\theta_1, \theta_2, \dots$) に対して，いつでも尤度 $L(\theta)$ ($L(\theta_1, \dots)$) が考えられる．
標本が与えられたときの母数の推定量として，$L(\theta)$ を最大にする $\hat{\theta}$ を考えることができる．

$$\hat{\theta} = \mathrm{argmax}_{\theta} L(\theta)$$

これを最尤推定 (MLE=Maximum Likelyhood Estimation), $\hat{\theta}$ を最尤推定量という．
$\frac{\mathrm{d}}{\mathrm{d}\theta} L(\theta) = 0$ を解けばよいが，対数尤度 $\log L(\theta)$ に対して $\frac{\mathrm{d}}{\mathrm{d}\theta} \log L(\theta) = 0$ を考えるほうが楽なことが多い．

L02-Q1

Quiz(ベルヌイ分布の母数の最尤推定)

$X \sim \mathrm{B}(1, p)$ のサイズ N の標本として，$x_1, \ldots, x_N$ が得られた．ただし，$\sum_{i=1}^{n} x_i = k \leq N$ だった．
母数 p を最尤推定しよう．

L02-Q2

Quiz(一様分布の母数の最尤推定)

$X \sim \mathrm{U}(c, d)$ のサイズ N の標本として，$x_1, \dots, x_N$ が得られた．

1. 母平均値，母分散を推定しよう．これらの値となる母数 c, d を求めよう．
2. 母数 c, d を最尤推定しよう．

ここまで来たよ

1. 母比率の推定再訪・離散-連続型確率変数
2. 母数の最尤推定・ベイズ推定
   - 従来の母比率の点推定
   - 尤度と最尤推定
   - ベイズ推定

ベイズの定理 I

定理 (ベイズの定理 岩薩林 (2.18),(2.19))

$$p_{X|Y}(x|y) \overset{\text{定義}}{=} \frac{p(x,y)}{p_Y(y)} = \frac{p_{Y|X}(y|x)p_X(x)}{\sum_{x'} p_{Y|X}(y|x')p_X(x')},$$

$$p_{Y|X}(y|x) \overset{\text{定義}}{=} \frac{p(x,y)}{p_X(x)} = \frac{p_{X|Y}(x|y)p_Y(y)}{\sum_{y'} p_{X|Y}(x|y')p_Y(y')}.$$

$p_{X|Y}(x|y)$ を $p_{Y|X}(y|x')$ (と $p_X(x')$) で書き表す式,
$p_{Y|X}(y|x)$ を $p_{X|Y}(x|y')$ (と $p_Y(y')$) で書き表す式.
定義の分子分母を, 乗法公式と全確率の公式で書き直しただけ (なので余分の記憶は不要).

ベイズ推定

次のような確率変数 X_i, Y の同時分布 $f(x_i, y)$ を考えよう．
(出題者が指定する代わりに仏が)$Y \sim \mathrm{U}(0,1)$ から，実現値 $Y = y$ を選ぶ．$f_{\mathrm{Y}}(y) = u(y; 0, 1)$.
y が選ばれたとき，$X_i \sim \mathrm{B}(1, y)$ で，サイズ N の標本 $x_1, \ldots, x_N$ が抽出される．条件付き分布 $f_{\mathrm{X|Y}}(x|y) = y^x (1-y)^{1-x}$.
このとき，標本 $x_1, x_2, \ldots$ から y を推定する．
標本を条件とした条件付き分布 $f_{\mathrm{Y|X}}(y|x_1, \ldots, x_N)$ が得られる．

$$\boxed{\text{事後確率分布 } f_{\mathrm{Y|X}}(y|x)} \longleftarrow \boxed{\text{事前確率分布 } f_{\mathrm{Y}}(y)}$$

$$\uparrow$$

$$\boxed{\text{情報 } X = x}$$

追加の情報 x_{N+1} が得られるたびに, 事後確率分布は正確になっていく.

主観確率=信念 berief の考え方

上では，$\mathrm{U}(0,1)$ に従うという仏の行動を明示したが，仏は推定者の心の中のみにあると考えてもよい．
すなわち，事前確率分布 f_{Y} は，仏がどんな y を選んだかに対する，推定者の主観的な予想であってよい．
このとき，事後確率分布 $f_{\mathsf{y}|x}(y|x_1,\dots,x_N)$ は，標本という追加情報を受け取った上での，推定者の主観的な予想である．
事後確率分布は，y の確率を定めているので，母数についての確率 $P(c \leq Y \leq d)$ を考えてよい．
$\leftrightarrow$ 頻度主義では，この確率は 0 または 1.

ベイズ推定のインプットは，‘モデル’ or ‘尤度’ $f_{\mathrm{X}|\mathrm{Y}}(x|y)$, 事前確率分布 $f_{\mathrm{Y}}(y)$, 標本 $x_1, \ldots, x_N$.
ベイズ推定のアウトプットは，事後確率分布 $f_{\mathrm{Y}|\mathrm{X}}(y|x_1, \ldots, x_N)$.

- (点推定のように)1 個の代表値に集約する
- （四分位範囲や区間推定のように）1 個の代表的な部分集合に集約する

には複数の方法がありうる．（平均値，中間値，⋯）
1 個の代表値に集約する方法のひとつ

ベイズ MAP 推定　MAP，Maximum a posteriori，最大事後確率
事後確率（密度）を最大にする y を選ぶ：
$\mathrm{argmax}_y f_{\mathrm{Y}|\mathrm{X}}(y|x_1, \ldots, x_N)$

尤度でなく確率 (密度) なので，確率最大の値を選ぶ，と言える．

L02-Q3

Quiz(一様事前分布のもとでのベルヌイ分布の母数のベイズ推定)

$X \sim \mathrm{B}(1, y)$ の母数 y の事前分布を $Y \sim \mathrm{U}(0, 1)$ とする．
サイズ N の標本として，$x_1, \ldots, x_N$ が得られた．ただし，
$\sum_{i=1}^{n} x_i = k \leq N$ だった．
母数 y の事後分布を求めよう．
事後分布の確率密度関数が最大となる y を求めよう．

L02-Q4

Quiz(ベルヌイ分布のベイズ推定: 非一様事前分布)

$X \sim \mathrm{B}(1, y)$ の母数 y の事前分布を $f_{\mathrm{Y}}(y) = 6y(1-y)u(y; 0, 1)$ とする．
サイズ N の標本として，$x_1, \ldots, x_N$ が得られた．ただし，
$\sum_{i=1}^{n} x_i = k \leq N$ だった．
母数 y の事後分布を求めよう．
事後分布の確率密度関数が最大となる y を求めよう．