共役事前分布・母数の区間推定

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 L03(2026-04-27 Mon)

最終更新: Time-stamp: "2026-04-27 Mon 18:26 JST hig"

今日の目標

- 共役事前分布を説明できる
- 標本平均値，MLE, MAP, EAP, MED を説明できる
- 信頼区間，ベイズ信用区間，HPD=HDI を説明できる

L02-Q1

Quiz 解答: ベルヌイ分布の母数の最尤推定

尤度は $L(p) = \prod_{i=1}^{N} p^{x_i}(1-p)^{1-x_i} = p^k(1-p)^{N-k}$.
対数尤度は $\log L(p) = k\log p + (N-k)\log(1-p)$.
p の最尤推定値は,

$$0 = \frac{\mathrm{d}L}{\mathrm{d}p}(p) = \frac{k}{p} + \frac{N-k}{1-p}$$

を解いて, $p = \frac{k}{N}$.

L02-Q2

L02-Q3

Quiz 解答: 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

$$
\begin{aligned}
f_{\mathrm{Y|X}}(y|x_i) =& \frac{f_{\mathrm{X|Y}}(x_1, \dots, x_N|y) f_{\mathrm{Y}}(y)}{\int f_{\mathrm{X|Y}}(x|y) f_{\mathrm{Y}}(y)\,\mathrm{d}y} \\
\propto& \prod_i y^{x_i}(1-y)^{1-x_i} u(y;0,1) = y^k (1-y)^{N-k} u(y;0,1)
\end{aligned}
$$

事後分布の確率密度関数が最大となるのは，最尤推定 MLE と同じ値で，$y = \frac{k}{N}$.

L02-Q4

Quiz 解答: ベルヌイ分布のベイズ推定: 非一様事前分布

$$
\begin{aligned}
f_{\mathrm{Y|X}}(y|x_i) =& \frac{f_{\mathsf{x|y}}(x_1, \dots, x_N|y) f_{\mathrm{Y}}(y)}{\int f_{\mathsf{x|y}}(x|y) f_{\mathrm{Y}}(y)\,\mathrm{d}y} \\
\propto& \prod_i y^{x_i}(1-y)^{1-x_i} 6y(1-y) u(y;0,1) \\
\propto& y^{k+1}(1-y)^{N-k+1} u(y;0,1)
\end{aligned}
$$

事後分布の確率密度関数が最大となるのは，$y = \frac{k+1}{N+2}$.

ここまで来たよ

2 母数の最尤推定・ベイズ推定

3 共役事前分布・母数の区間推定
- 共役事前分布
- ベイズ事後分布に基づく様々な推定量

ベータ分布 (Beta distribution)

定義 (ベータ分布 赤石 p.43,p.86**)**

連続型確率変数 Y の確率密度関数が

$$f(y) = \begin{cases} \frac{1}{B(a,b)} y^{a-1}(1-y)^{b-1} & (0 \leq y \leq 1) \\ 0 & (\text{他}) \end{cases}$$

のとき，y はパラメタ a, b のベータ分布 $\mathrm{beta}(a,b)$ にしたがうという．ただし，全確率 1 の条件から定数 $B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$，　$\Gamma(n+1) = n!$.

- 実は，正の実数の a, b に対して定義される．
- $B(a,b), \Gamma(a)$ は，一般にベータ関数，ガンマ関数とよばれる特殊関数

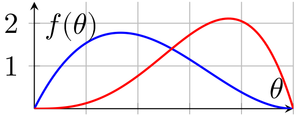
すでにベータ分布とは出会っていた！

定理 (ベータ分布の最頻値，母平均値)

$Y \sim \mathrm{beta}(a,b)$ のとき，$\mathrm{E}[Y] = \frac{a}{a+b}$, 最頻値 $= \frac{a-1}{a+b-2}$.

直観的に納得できる？ a, b の意味は⋯

- 一様な事前分布 $Y \sim \mathrm{U}(0,1)$ は，$Y \sim \mathrm{beta}(1,1)$ と同じ．
- サンプル $X_1, \ldots, X_N \sim \mathrm{B}(1,y)$ に対して，
  一様な事前確率密度関数 $Y \sim \mathrm{beta}(1,1)$ のとき，事後分布は，ベータ分布 $\mathrm{beta}(\sum_i x_i + 1, N - \sum_i x_i + 1)$.
- 事前分布 $Y \sim \mathrm{beta}(a,b)$ のとき，事後分布は，ベータ分布
  $\mathrm{beta}(\sum_i x_i + a, N - \sum_i x_i + b)$.



L03-Q1

Quiz(ベータ分布)

$\Theta \sim \mathrm{beta}(a, b)$ のとき，すなわち，確率密度関数が

$$f(\theta) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta^{a-1}(1-\theta)^{b-1} \quad (0 \le \theta \le 1)$$

のとき，次を示そう．

1. $\mathrm{E}[\Theta] = \frac{a}{a+b}$.
2. $\mathrm{V}[\Theta] = \frac{ab}{(a+b)^2(a+b+1)}$.
3. Θ の最頻値は $\frac{a-1}{(a-1)+(b-1)}$.

共役性

定義 (共役性)

Y の分布のクラス（例：ベータ分布，パラメタさまざま）が確率モデル $f_{\mathrm{X}|\mathrm{Y}}(x|y)$ に関して共役であるとは，
事前分布がそのクラスに属するなら事後分布もそのクラスに属する，ことをいう．

一般にベイズ推定すると，事前分布 (関数) は別の形の事後分布 (関数) に変化する

共役な事前分布のときは，パラメタだけ変えた，同じクラスの事後分布 (関数) に変化する
更新：$(a, b) \mapsto (a + k, b + N - k)$.
事前がベータなら，何回ベイズ更新しても事後はベータのまま．分布の空間の中の，この変換による不変部分空間になっている．
この枠内でやるなら，計算機実験不要，a, b の更新式さえあればいい．

共役でない事前分布の例

$$f_{\mathrm{Y}}(y) = \frac{2}{\pi} |\sin(c\pi y)|$$

確率モデルとその共役事前分布

	確率モデルとパラメタ	共役事前分布
L02,L03	ベルヌイ分布の p	ベータ分布
	二項分布の p	ベータ分布
	多項分布の $p_1, \dots, p_k$	ディリクレ分布
	ポアソン分布の母平均値	ガンマ分布
	ガンマ分布のパラメタ	ガンマ分布
L04	正規分布の母平均値	正規分布
	正規分布の母分散	逆ガンマ分布
	多変量正規分布の母共分散行列	逆ウィシャート分布

事前分布の選び方

とりあえず理論的な振る舞いを知りたいとき

- 共役事前分布 (教科書や Web を調べる)

このへんに分布するだろう，というドメイン知識がある場合

- そのあたりにピークを持つ，だけど広い分布

何の事前情報もない場合，客観性を主張する必要がある場合

- ほぼ一定で，分散が大きい分布例 $\mathrm{U}(0, 100), \mathrm{N}(0, 100^2)$.
- 負の値はとらない，などの情報があるなら，それを反映させることもある例 カイ二乗分布，半正規分布
- 全確率 $= +\infty$ となるかのような「分布」（improper prior という，例：$f_Y(y) = 1$）をとって，ベイズの公式の分子分母で $\infty/\infty = 1$ としてしまっても実害はない（MLE に相当）．

ここまで来たよ

2 母数の最尤推定・ベイズ推定

3 共役事前分布・母数の区間推定
- 共役事前分布
- ベイズ事後分布に基づく様々な推定量

ベイズ推定のアウトプットは Y の事後分布 $f_{\mathrm{Y}|\mathrm{X}}(y|x_1,\ldots,x_N)$.

- (標本平均値のように)1 個の代表値に集約
- (四分位範囲や信頼区間のように) 1 個の代表的な部分集合に集約

1 個の代表値に集約する方法

ベイズ MAP 推定 Maximum a posteriori (母最頻値)$\mathrm{argmax}_y f_{\mathrm{Y}|\mathrm{X}}(y|x_1,\ldots,x_N)$

ベイズ EAP 推定 Expected a posteriori (母平均値)
$$\mathrm{E}[Y|x_1,\ldots,x_N]=\int_{-\infty}^{+\infty} y\cdot f_{\mathrm{Y}|\mathrm{X}}(y|x_1,\ldots,x_N)\,\mathrm{d}y.$$

ベイズ MED 推定 $\mathrm{Median}[\mathrm{Y}]$ (母中央値)
s.t. $\int_{-\infty}^{\mathrm{Median}[Y]} f_{\mathrm{Y}|\mathrm{X}}(y|x_1,\ldots,x_N)\,\mathrm{d}y=\frac{1}{2}$.

cf. ベイズでない方法：MLE, 標本平均値（母平均値が母数の場合）

1 個の代表的な部分集合に集約する方法

信頼区間，ベイズ信用区間，HPD=HDI

L03-Q2

Quiz(区分的線形な事後確率密度のベイズ MAP, EAP, MED 推定)

ベイズ推定の結果，パラメタ y の事後確率密度が次の分布で与えられた．

$$f(y) = \begin{cases} \frac{1}{5}(1 - |y+1|) & (-2 < y \leq 0) \\ \frac{4}{5}(1 - |y-1|) & (0 < y \leq 2) \\ 0 & (\text{他}) \end{cases}$$

1. Y をベイズ MAP 推定しよう．
2. Y をベイズ EAP 推定しよう．
3. Y をベイズ MED 推定しよう．

信頼区間 (Confidence Interval)

定義 (頻度論的な信頼区間)

母数 θ の区間で，上下限が確率変数であるようなもの $[\ell(X_1, \dots, X_N), u(X_1, \dots, X_N)]$ が，パラメタ θ に対する，信頼係数 $1-\alpha$ の信頼区間であるとは，

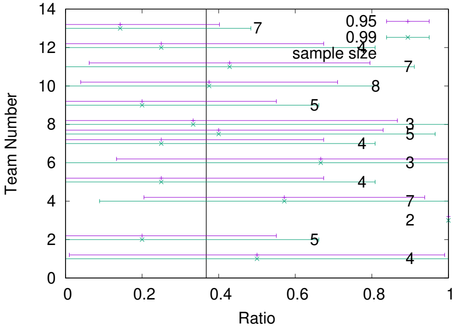
$$P(\ell(X_1, \dots) < \theta < u(X_1, \dots)) = 1 - \alpha$$

であることをいう．

これに，実現値 $X_i = x_i$ を代入したものも信頼区間とよばれる．
しばしば，$P(\theta < \ell(X_1, \dots)) = P(\theta > u(X_1, \dots)) = \frac{\alpha}{2}$ であるように（上下対称であるように）とる．

信頼区間の意味

頻度主義的な概念．
母数 θ は未知だが定数，標本抽出のたびに実現値が得られる区間の両端が確率変数．
信頼係数 $1-\alpha$ の信頼区間とは，標本抽出を繰り返したとき,）θ がその区間に含まれる確率が $1-\alpha$.



L03-Q3

Quiz(母比率の区間推定)

選挙で出口調査をしたところ, 50 人中 35 人が A 候補に投票したと答えた. 母集団を投票した人 2500 人とする. そのうち A 候補に投票した人の母比率 (得票率) を考える.

1. A 候補の得票率を, (点) 推定しよう
2. A 候補の得票率を, 信頼係数 $1-\alpha=0.95$ で区間推定しよう.
3. A 候補の得票率を, 信頼係数 $1-\alpha=0.99$ で区間推定しよう.

ベルヌイ分布の母数 p の区間推定

中心極限定理による近似

$$P[\theta + \Phi^{-1}(\frac{\alpha}{2})\sqrt{\frac{\theta(1-\theta)}{N}} < \overline{X} < \theta + \Phi^{-1}(1-\frac{\alpha}{2})\sqrt{\frac{\theta(1-\theta)}{N}}] = 1-\alpha$$

移項

$$P[\overline{X} - \Phi^{-1}(1-\frac{\alpha}{2})\sqrt{\frac{\theta(1-\theta)}{N}} < \theta < \overline{X} - \Phi^{-1}(\frac{\alpha}{2})\sqrt{\frac{\theta(1-\theta)}{N}}] = 1-\alpha$$

近似 $\theta = \overline{X}$

$$P[\overline{X} - \Phi^{-1}(1-\frac{\alpha}{2})\sqrt{\frac{\overline{X}(1-\overline{X})}{N}} < \theta < \overline{X} - \Phi^{-1}(\frac{\alpha}{2})\sqrt{\frac{\overline{X}(1-\overline{X})}{N}}] = 1-\alpha$$

事後分布からの区間推定

- ベイズ信用区間＝ベイズ確信区間 credible interval
- 最高事後密度領域＝HPD 領域＝最高密度区間＝HDI

いずれも，「（確率変数である）母数 θ が，この区間に含まれる確率は $1-\alpha$」と言ってよい．

信用区間 (credible interval)

定義 (ベイズ信用区間 赤石 p.53**)**

観測データ $X = x$ に基づく Yの区間 $[\ell(x), u(x)]$ が，パラメタ Yに対する，信頼係数 $1-\alpha$ の信用区間であるとは，

$$P(\ell(x) < Y < u(x)) = 1-\alpha$$

が成り立つことをいう．

しばしば，左右両端の同じ確率だけ捨てる，すなわち，
$P(Y < \ell(x)) = P(Y > u(x)) = \frac{\alpha}{2}$ であるようにとる．
y_q を $Y|X = x$ の q 分位数とすると，$\ell(x), u(x)$ は，$\frac{\alpha}{2}, 1-\frac{\alpha}{2}$ 分位数．
$\rightsquigarrow$ Colab

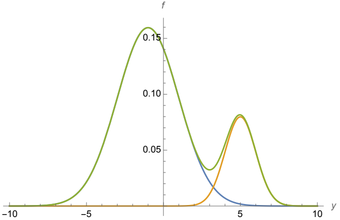
最高事後密度 (HPD=highest posterior density) 領域

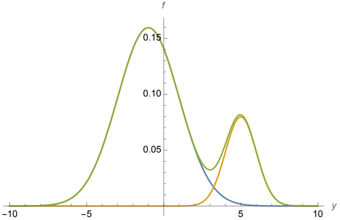
定義 (最高事後密度 (HPD) 領域 赤石 p.53)

観測データ $X=x$ に基づく $(1-\alpha)$ 最高事後密度領域とは，次を満たすパラメタ空間の部分集合 $s(x)$ のこと．

$$P(\theta \in s(x)|X=x) = 1-\alpha$$
$$(y_a \in s(x) \text{かつ} y_b \notin s(x)) \Rightarrow f_{\mathrm{Y|X}}(y_a|x) > f_{\mathrm{Y|X}}(y_b|x)$$

ぎりぎりの密度 $f_c \in \mathbb{R}$ で，$s(y) = \{y|f(y|x) \geq f_c\}$ の形で書かれるはず．
HDI=Highest Density Interval





L03-Q4

Quiz(区分的線形な事後確率密度の信用区間と最高事後密度領域)

ベイズ推定の結果，パラメタ y の事後確率密度が次の分布で与えられた．

$$f(y) = \begin{cases} \frac{1}{5}(1-|y+1|) & (-2 < y \leq 0) \\ \frac{4}{5}(1-|y-1|) & (0 < y \leq 2) \\ 0 & (\text{他}) \end{cases}$$

1. Y の（左右対称の）$1-\alpha = 0.95$ ベイズ信用区間=CI を求めよう．
2. Y の $1-\alpha = 0.95$ 最高事後密度領域=HPD 領域=HDI を求めよう．

事後予測分布

事後分布は母数 y の分布だが，事後予測分布は，母数の事後分布に基づく，次の測定値 $\tilde{x}$ の分布を，全確率の法則（同時分布→周辺分布）で求めたもの．

$$f(\tilde{x}) = \int_{-\infty}^{+\infty} f_{\mathrm{X|Y}}(\tilde{x}|y) f_{\mathrm{Y|X}}(y|x_1, \dots, x_N)\, \mathrm{d}y$$

この分布も，単一の値（母最頻値，⋯）や，（X がはいる確率が p の）区間に集約できる．

L03-Q5

Quiz(ベルヌイ分布の事後予測分布)

確率変数 X は, パラメタ y のベルヌイ分布 $\mathrm{B}(1, y)$ にしたがう.
Y の事前分布 を $\mathrm{beta}(1, 1)$ とする.
確率変数 X の標本 $x_1, \ldots, x_N$ を得て, $\sum_i x_i = k$ だった.

1. ベイズ事後分布を求めよう.
2. ベイズ事後予測分布を求めよう.