

事後予測分布・多パラメタのベイズ推定・一般化線形モデル

樋口さぶろお <https://hig3.net>

龍谷大学大学院 先端理工学部研究科 数理・情報科学コース

計算科学特論 L07(2024-11-06 Wed)

最終更新: Time-stamp: "2024-11-06 Wed 13:55 JST hig"

今日の目標

- 事後予測分布を説明できる
- 多母数のベイズ推定を説明できる
- 一般化線形モデルを説明できる



L06-Q3

Quiz 解答: ベイズ推定

- ① 一般に, $\Theta|Y \sim \text{beta}(a, b)$ ($a, b > 1$) のとき, $f(\theta|y)' = 0$ となるのは $\theta = \frac{a-1}{a+b-2}$. よって MAP 推定値は, $\theta = \frac{2-1}{2+3-2} = \frac{1}{3} = 0.3333$.
- ② 一般に, $\Theta|Y \sim \text{beta}(a, b)$ ($a, b > 1$) のとき, $E[\Theta|Y] = \frac{a}{a+b}$. よって EAP 推定値は, $\theta = \frac{2}{2+3} = \frac{2}{5} = 0.4$.
- ③ MED 推定値は, `stats.beta(2,3).ppf(0.5)=0.3857`.

ここまで来たよ

6 正規モデルのベイズ推定

7 事後予測分布・多パラメタのベイズ推定・一般化線形モデル

- 事後予測分布
- 正規モデル（母平均値，母分散パラメタ）
- 一般化線形モデル

事後予測分布

予測分布とは、確率モデル $f(y|\theta)$, 標本 $y_1, \dots, y_N$ を元に推測する, 次の観測値 $\tilde{Y}|Y$ のしたがる分布のこと.

頻度論例えば, $Y \sim N(\theta, \sigma^2)$, θ 未知, σ^2 既知のとき,

まず, θ の点推定 $\hat{\theta} = \bar{y} = \frac{1}{N}(y_1 + \dots + y_N)$. 予測分布 $\tilde{Y}|Y \sim N(\hat{\theta}, \sigma^2)$.

ベイズ的一般論で,

確率モデル $f(y|\theta)$ を, θ の事後分布 $f_{\theta|Y}(\theta|y)$ で重み付けして加えたものを**事後予測分布**という.

$$\begin{aligned} f_{Y|Y:}(\tilde{y}|y_1, \dots, y_N) &= \int_{-\infty}^{+\infty} f_{Y, \theta|Y:}(\tilde{y}, \theta|y_1, \dots, y_N) d\theta \\ &= \int_{-\infty}^{+\infty} f_{Y|\theta, Y:}(\tilde{y}|\theta, y_1, \dots, y_N) f_{\theta|Y}(\theta|y_1, \dots, y_N) d\theta \\ &= \int_{-\infty}^{+\infty} f_{Y|\theta}(\tilde{y}|\theta) f_{\theta|Y}(\theta|y_1, \dots, y_N) d\theta. \end{aligned}$$

事後予測分布（二項モデル）

Hoff ベイズ p.45

事前分布 $\Theta \sim \text{beta}(a, b)$, 観測値 $y_1, \dots, y_N$ のもとでの.

$$\begin{aligned} & P(\tilde{Y} = 1 | Y_1 = y_1, \dots) \\ &= \int_0^{+\infty} P(\tilde{Y} = 1 | \Theta = \theta) \times f(\theta | y_1, \dots, y_N) d\theta \\ &= \int_0^{+\infty} \theta \times f(\theta | y_1, \dots, y_N) d\theta \\ &= E[\Theta | y_1, \dots] = \frac{a + \sum_i y_i}{a + b + N} \end{aligned}$$

事後予測分布 (ポアソンモデル)

Hoff ベイズ p.54

事前分布 $\Theta \sim \text{gamma}(a, b)$, 観測値 $y_1, \dots, y_N$ のもとでの.

$$\begin{aligned} f_{Y|Y:}(\tilde{y}|y_1, \dots, y_N) &= \int_0^{+\infty} f(\tilde{y}|\theta) \times f(\theta|y_1, \dots, y_N) d\theta \\ &= \int_0^{+\infty} \frac{1}{\tilde{y}!} \theta^{\tilde{y}} e^{-\theta} \times \frac{(b+N)^{a+\sum_i y_i}}{\Gamma(a+\sum_i y_i)} \theta^{a+\sum_i y_i-1} e^{-(b+N)\theta} d\theta \\ &= \dots = \text{'パラメタ } a + \sum_i y_i, b+N \text{ の負の二項分布'} \end{aligned}$$

$\tilde{Y}$ 離散

Θ 連続

事後予測分布（正規モデル（分散固定））

Hoff ベイズ p.81

事前分布 $\theta = \mu \sim N(\mu_0, \tau_0^2)$, 観測値 $y_1, \dots, y_N$ のもとでの.

$$f_{Y|Y}(\tilde{y}|y_1, \dots, y_N) = \int_{-\infty}^{+\infty} f(\tilde{y}|\theta) \times f(\theta|y_1, \dots, y_N) d\theta$$

正規分布にしたがう独立な確率変数 $X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N(\mu_2, \sigma_2^2)$ に対して, 和 $S = X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ である.

$$\Theta \sim N\left(\frac{\tau_0^{-2}\mu_0 + N\sigma^{-2}\frac{1}{N}\sum_i y_i}{\tau_0^{-2} + N\sigma^{-2}}, (\tau_0^{-2} + N\sigma^{-2})^{-1}\right).$$

$$\tilde{Y} \sim N(\Theta, \sigma^2) \text{ より,}$$

$$\tilde{Y} = \epsilon + \Theta.$$

$$\epsilon \sim N(0, \sigma^2).$$

$$\tilde{Y} \sim N\left(\frac{\tau_0^{-2}\mu_0 + N\sigma^{-2}\frac{1}{N}\sum_i y_i}{\tau_0^{-2} + N\sigma^{-2}} + 0, (\tau_0^{-2} + N\sigma^{-2})^{-1} + \sigma^2\right).$$

L07-Q1

Quiz(正規分布の母平均値のベイズ推定)

確率変数 Y は, 正規分布にしたがう. すなわち,

$$f(y|\mu) = \frac{1}{\sqrt{2\pi 2^2}} e^{-\frac{(y-\mu)^2}{2 \cdot 2^2}}.$$

母数 μ の事後分布が

$$f(\mu|y_1, \dots, y_N) = \frac{1}{\sqrt{2\pi 3^2}} e^{-\frac{(\mu-5)^2}{2 \cdot 3^2}}$$

であるとき, 次の測定値 $\tilde{y}$ の事後予測分布 $f(\tilde{y}|y_1, \dots, y_N)$ を求めよう.

ここまで来たよ

6 正規モデルのベイズ推定

7 事後予測分布・多パラメタのベイズ推定・一般化線形モデル

- 事後予測分布
- 正規モデル（母平均値，母分散パラメタ）
- 一般化線形モデル

正規モデル (母平均値, 母分散パラメタ)

Hoff ベイズ §5.3

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 y g のドーナツを作り出す.

ここまでの μ が (μ, σ^2) になると思えばよい.

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

2 変数の事後分布

$$\begin{aligned} f(\mu, \sigma^2|y_1, \dots, y_N) &= C f(y_1, \dots, y_N|\mu, \sigma^2) f(\mu, \sigma^2) \\ &= C f(y_1, \dots, y_N) f(\mu|\sigma^2) f(\sigma^2) \end{aligned}$$

共役分布を探そう. μ については正規分布. $\sigma^2 > 0$ については知ってる分布でないが, $\tilde{\sigma}^2 = 1/\sigma^2 > 0$ を変数と見直すと, ガンマ分布が共役分布になりそう.

この確率モデル $f(y|\mu, 1/\tilde{\sigma}^2) = \frac{(\tilde{\sigma}^2)^{1/2}}{\sqrt{2\pi}} e^{-\frac{1}{2}\tilde{\sigma}^2(y-\mu)^2}$

ガンマ分布 $f(\theta|a, b) = \frac{b^a}{\Gamma(a)} \theta^{a-1} e^{-b\theta} \quad (\theta > 0)$

$\tilde{\sigma}^2 \sim \text{gamma}(a, b)$ のとき, 逆数 $\sigma^2 = 1/\tilde{\sigma}^2$ のしたがう分布を, **逆ガンマ分布** (inverse gammadistribution) という.
 $\sigma^2 = 1/\tilde{\sigma}^2 \sim \text{inverse-gamma}(a, b)$.

L07-Q2

Quiz(確率変数の変数変換)

一様分布にしたがう確率変数 $X \sim U(1, 3)$ に対して, $Y = 1/X$ を考える. Y の確率密度関数を求めよう.

L07-Q3

Quiz(確率変数の変数変換 (対数正規分布))

正規分布にしたがう確率変数 $X \sim N(\mu, \sigma^2)$ に対して, $Y = e^X$ を考える. Y は対数正規分布にしたがうという (対数をとったら正規分布にしたがう). Y の確率密度関数を求めよう. グラフの概形を予想しよう.

事前分布

$$\tilde{\sigma}^2 = 1/\sigma^2 \sim \text{gamma}(a, b)$$

$$\mu|\sigma \sim N(\mu_0, \tau_0^2)$$

$$Y_1, \dots, Y_n | \Theta, \sigma^2 \sim N(\mu, \sigma^2)$$

さらに, τ_0^2 が σ^2 に正比例するように修正する: $\tau_0^2 = \sigma^2 / \kappa_0$.

a, b は, あとで解釈しやすくなるように, $a = \frac{1}{2}\nu_0, b = \frac{1}{2}\nu_0\sigma_0^2$ とおく (パラメトライズする).

$\kappa_0, \nu_0, \sigma_0$ が確率分布を持たない本当の定数.

事後分布

μ の事後分布は正規分布に戻る. $\mu|y_1, \dots, y_N, \sigma^2 \sim N(\frac{\kappa_0\mu_0 + N\bar{y}}{\kappa_0 + N}, \frac{\sigma^2}{\kappa_0 + N})$.

σ^2 の事後分布はガンマ分布に戻る.

$$\tilde{\sigma}^2|y_1, \dots, y_N \sim \text{gamma}(\frac{\nu_0 + N}{2}, \frac{\nu_0 + N}{2} \sigma_N^2),$$

$$\sigma_N^2 = \frac{1}{\nu_0 + N} [\nu_0 \sigma_0^2 + \sum_{i=1}^N (y_i - \bar{y})^2 + \frac{\kappa_0 N}{\kappa_0 + N} (\bar{y} - \mu_0)^2].$$

$$\bar{y} = \frac{1}{N} \sum_i y_i.$$

情報の統合 略

事後予測分布 略

ここまで来たよ

6 正規モデルのベイズ推定

7 事後予測分布・多パラメタのベイズ推定・一般化線形モデル

- 事後予測分布
- 正規モデル（母平均値，母分散パラメタ）
- 一般化線形モデル

(復習) 線形回帰モデル

$$Y_i = \beta_1 + \beta_2 x_i + \epsilon_i$$

- 説明変数 x_i : 例: 加熱前の棒の長さ (人間が決められる変数)
- 目的変数 Y_i : 例: 加熱後の棒の長さ (確率変数)
- ノイズ ϵ_i (確率変数) $\sim N(0, \sigma^2)$, 独立同分布.
- β_1, β_2 回帰係数 (測定番号 $i = 1, \dots, N$)

x_i, Y_i は実数値をとる (本当は正の)

久保拓也, <https://www.iwanami.co.jp/book/b257893.html>, 岩波書店 (2012).

目的変数が自然数に値をとる回帰モデル？

植物の高さ x_i から、(今シーズンの) 種子の個数 Y_i を予測したい.
線形回帰モデル？

$$Y_i = \beta_1 + \beta_2 x_i + \epsilon_i?$$

♡ x_i が大きいと Y_i も大きそう.

♠ 正規分布ではないはず.

♠ 高さ x_i が、同じなら、本来、種子の個数はポアソン分布にしたがうはず.

♠ 高さ x_i は実数、計算された Y_i も実数になっちゃう？

♠ 個数が負になる心配.

一般化線形モデル (GLM) としてのポアソン回帰モデル

意味や $\lambda > 0$ を考えると、次が適切.

ポアソン回帰モデル

$$Y_i \sim \text{Po}(\lambda_i)$$
$$\log \lambda_i = \beta_1 + \beta_2 x_i$$

次の3つを選んで作れるものを、一般化線形モデル (GLM=Generalized Linear Model) という.

- 確率分布 $\text{Po}(\lambda)$ パラメタを説明変数で説明する
- リンク関数 $g(\lambda) = \log(\lambda)$ 1 次関数に等置するための関数
- 線形予測子 $\beta_1 + \beta_2 x_i$ 1 次式. 重回帰のように変数が増えることがありうる. $\beta_1 + \beta_2 x_i + \beta_3 x_i^2$ も線形と考える.

一般化線形モデル (GLM) としての線形回帰モデル

線形回帰モデルも一般化線形モデルの例

$$Y_i \sim N(\mu_i, \sigma^2)$$

$$\mu_i = \beta_1 + \beta_2 x_i$$

- 確率分布 $N(\mu, \sigma^2)$ パラメタを説明変数で説明する
- リンク関数 $g(\mu) = \mu$ 1 次関数に等置するための関数
- 線形予測子 $\beta_1 + \beta_2 x_i$ 1 次式. 重回帰のように変数が増えることはありうる

一般化線形モデル (GLM) としてのロジスティック回帰

ロジスティック回帰も一般化線形モデルの例

画像 i のある特徴量 x_i から、画像が猫である確率 $0 \leq p_i \leq 1$ を計算する

$$p_i = \frac{e^{\beta_1 + \beta_2 x_i}}{1 + e^{\beta_1 + \beta_2 x_i}} = \frac{1}{e^{-(\beta_1 + \beta_2 x_i)} + 1}.$$

ロジスティック関数 $p = \text{logistic}(z) = \frac{1}{e^{-z} + 1}$ の逆関数はロジット関数 $z = \text{logit}(p) = \log \frac{p}{1-p}$ (ログオッズ)。

$$Y_i \sim B(1, p_i)$$

$$\text{logit} p_i = \beta_1 + \beta_2 x_i$$

- 確率分布 $B(1, p)$ パラメタを説明変数で説明する
- リンク関数 $g(p) = \text{logit}(p)$ 1 次関数に等置するための関数
- 線形予測子 $\beta_1 + \beta_2 x_i$ 1 次式. 重回帰のように変数が増えることはありうる

一般化線形モデル (GLM) のパラメタ推定

たとえばポアソン回帰の場合、尤度は

$$L(\beta_1, \beta_2 | y_1, \dots, y_N) = \prod_{i=1}^N \frac{\lambda_i^{y_i}}{y_i!} e^{-\lambda_i}, \quad \lambda_i = \beta_1 + \beta_2 x_i$$

パラメタ β_1, β_2 を尤度最大化で決める (最尤推定).

対数尤度 $\log L(\beta_1, \beta_2 | y_1, \dots, y_N)$ のほうが、和になって考えやすいことが多い.