ベイズ統計

古谷知之

講義概要

- R演習：自然共役事前分布

 ①二項分布とベータ分布

 ②ポアソン分布とガンマ分布

 ③正規分布（平均）の推定

 ④正規分布（平均と分散）の推定

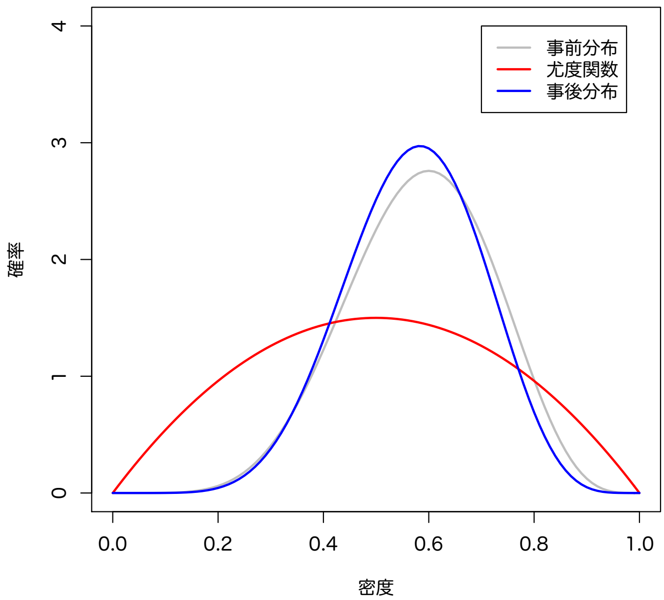
時間があれば…

- 尤度原理
- 継起の法則と理由不十分の原則
- 薬の効用とベイズ推定

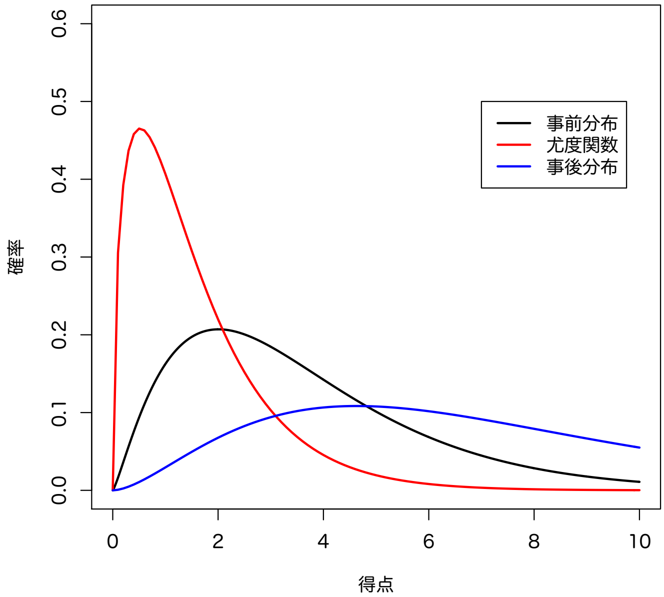
統計モデルの種類

	推定方法	確率分布	個体差/地域差
線形回帰モデル	最小二乗法	正規分布	なし
↓ 一般化線形モデル	最尤推定法	正規分布以外の分布も可能	なし
↓ 一般化線形混合モデル マルチレベルモデル	最尤推定法	正規分布以外の分布も可能	あり
↓ 階層ベイズモデル	ベイズ推定	自由に設定可能	

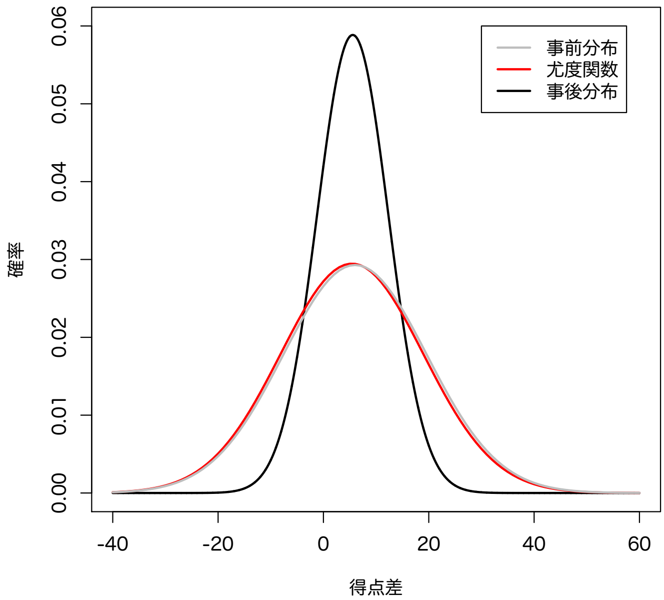
自然共役事前分布①
二項分布とベータ分布



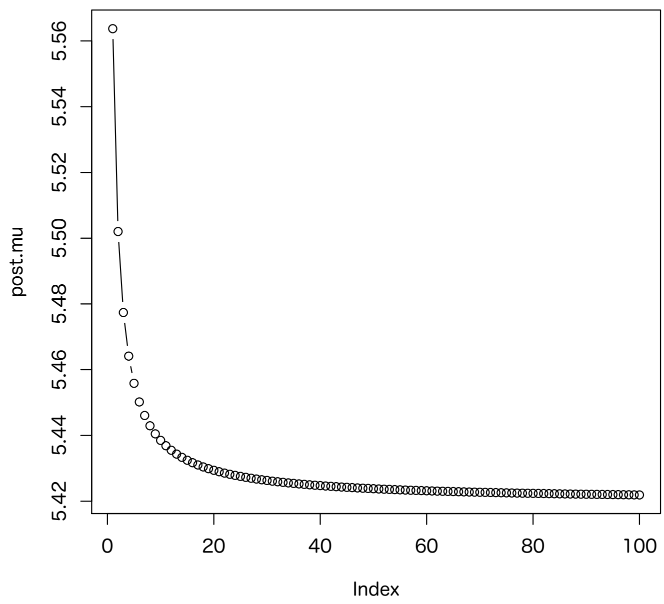
自然共役事前分布②
ポアソン分布とガンマ分布



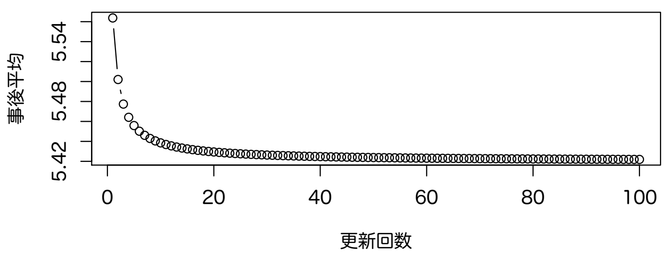
自然共役事前分布③
正規分布（平均）の推定

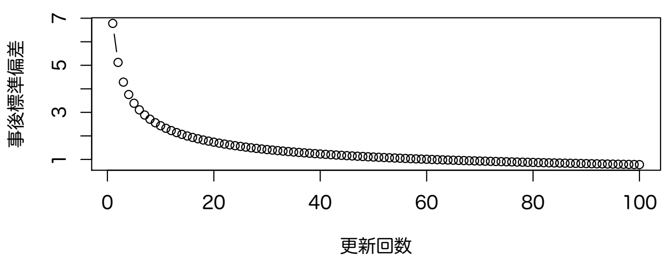


自然共役事前分布③
正規分布（平均）の推定



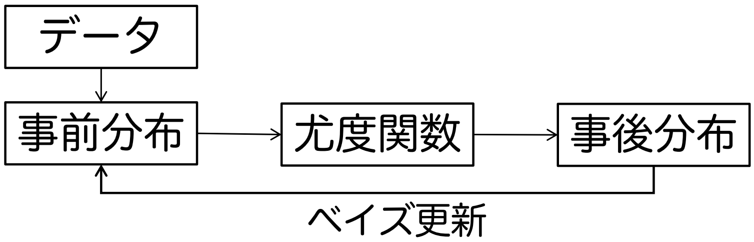
自然共役事前分布④
正規分布（平均と分散）の推定





経験ベイズ推定

- 事前分布をデータから与える方法を経験ベイズ(empirical Bayes)推定という
- 事前分布のパラメータを超パラメータ(hyper parameter)という
  - 事前分布の分散(精度)など



精度(precision)

- 正規分布をもちいてベイズ推定する際に、分散の逆数を精度として表すことがある

$$\tau = \frac{1}{\sigma^2}$$

- ベイズ統計では、精度がより重要である
- 分散が小さい⇔精度が高い
- 事後正規分布の精度は、事前正規分布の精度とデータの精度の和
- 事後正規分布の平均は、事前正規分布の精度とデータの精度の平均

正規分布の分散のベイズ推定

- 正規分布のベイズ推定における尤度関数

$$\prod_{i=1}^{n} f(x_i|\mu,\sigma^2) \propto \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(x_i-\mu)^2}{2\times\sigma^2}\right)$$

$$\propto \left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}} exp\left(-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\times\sigma^2}\right)$$

- ここで$\alpha-1=n/2, \lambda=\sum_{i=1}^{n}(x_i-\mu)^2/2$と置き換えると、

$$\prod_{i=1}^{n} f(x_i|\mu,\sigma^2) \propto \left(\frac{1}{\sigma^2}\right)^{\alpha-1} exp\left(-\frac{\lambda}{2\times\sigma^2}\right)$$

正規分布の分散のベイズ推定

- さらに、$\tau=\frac{1}{\sigma^2}$と置き換えると

$$\prod_{i=1}^{n} f(x_i|\mu,\sigma^2) \propto (\tau)^{\alpha-1} exp(-\lambda\cdot\tau)$$

- これはガンマ関数$Ga(\alpha,\lambda)$に比例する

$$Ga(\alpha,\lambda) \propto (\tau)^{\alpha-1} exp(-\lambda\cdot\tau)$$

- すなわち事前分布の精度τに対して、自然共役事前分布としてガンマ分布を用いることと同じ意味を持つ

ガンマ関数

- ガンマ関数$\Gamma(\alpha)$は次式で表される

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x}\, dx$$

$$\alpha > 0$$

- ガンマ関数$\Gamma(\alpha)$は以下のような性質を持つ

$$\Gamma(\alpha) = (\alpha - 1)!$$
$$\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$$
$$\Gamma(1/2) = \sqrt{\pi}$$

ガンマ分布

- ガンマ分布の確率密度関数$f(x)$は次式のようになる

$$f(x) = Ga(\alpha, \lambda) = \frac{\lambda^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, x \geq 0$$

- ガンマ分布の確率変数をXとするとき、平均$E(X)$と分散$V(X)$は以下の通り

$$E(X) = \frac{\alpha}{\lambda}$$

$$E(X) = \frac{\alpha}{\lambda^2}$$

正規分布の分散のベイズ推定

- さらにいえば、事前分布の分散σ^2に対する自然共役事前分布として、逆ガンマ分布を用いることと同じ意味である。
- 逆ガンマ分布は、正規分布のベイズ推定にはよく用いられる

- もっとも、事前分布の分散の事前分布として逆ガンマ分布を用いず、単に分散の逆数などとして与える場合もある=無情報事前分布

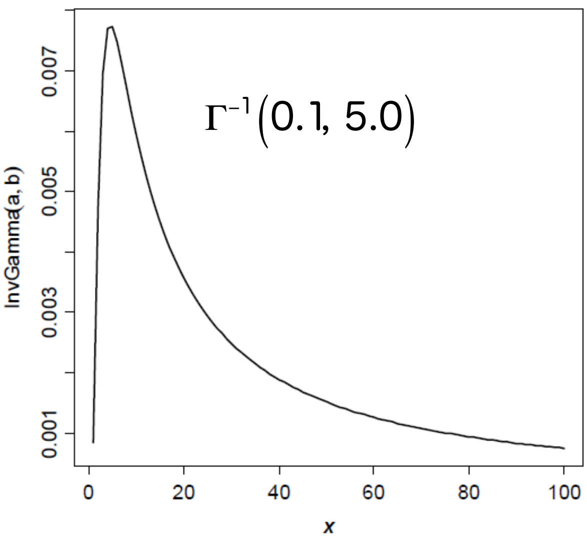
逆ガンマ分布

- 逆ガンマ分布

$$Ga^{-1}(\alpha, \lambda)$$

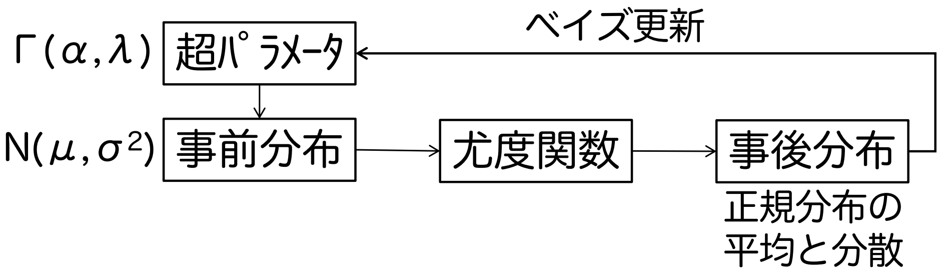
- 性質
  - 平均：$\lambda/(\alpha-1), \alpha > 1$
  - 分散：

$$\frac{\lambda^2}{(\alpha-1)^2(\lambda-1)}$$
$$\lambda > 1, \alpha > 2$$



階層ベイズ推定

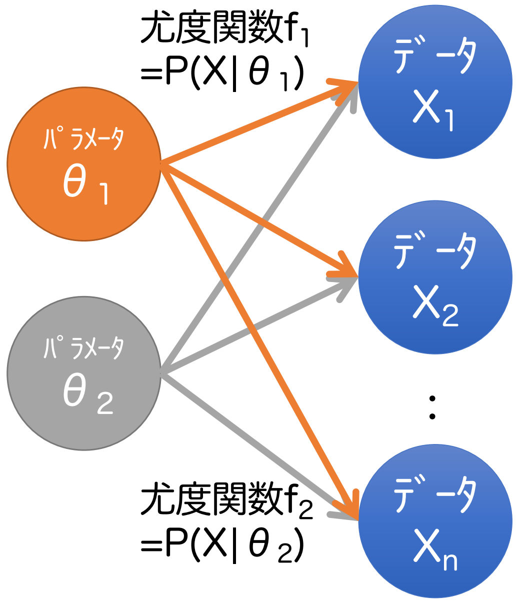
- 事前分布をベイズ推定する方法を階層ベイズ(hierarchical Bayes)推定という
- 事前分布の超パラメータをベイズ推定する
    - 事前正規分布の分散(逆ガンマ分布に従う)の超パラメータのベイズ推定



ベイズ推定を行う上での疑問

- 事前分布はどのような確率密度分布でもよいのか？
- ベイズ更新の際、事前分布はどの順番で与えても同じ事後分布が得られるのか？
- 同じ事前分布についてどのような尤度関数でも同じ事後分布が得られるのか？

尤度：パラメータからデータを再現する



尤度原理

- 未知パラメータθを推定するとき、尤度に観測されたデータの全ての情報を与えている場合には、比例関係にある尤度関数からは、同じ事前分布に対して同じ推論結果が導かれる

- ベイズ推定法は尤度原理を満たす
- 最尤推定法は尤度原理を満たさない

尤度原理

- コイントスをして10回中6回が表だった
- このとき尤度関数f_1は

$$f_1(x|\theta) = {}_{10}C_6\theta^6(1-\theta)^4 \propto \theta^6(1-\theta)^4$$

- 9回コインを投げて５回表が出た後に10回目に表が出る確率密度関数f_2は

$$f_2(x|\theta) = \binom{10-1}{6-1}\theta^{6-1}(1-\theta)^4 \propto \theta^6(1-\theta)^4$$

- この２つの関数は比例関係にある

$$f_1(x|\theta) \propto f_2(x|\theta)$$

尤度原理

- ちなみに頻度主義統計学においては、帰無仮説H_0と対立仮説H_1を次のようにおく

$$H_0: \theta = 0.5, H_1: \theta > 0.5$$

- 帰無仮説H_0 は有意水準5%で棄却される

$$P(X \geq 6|H_0) = \sum_{x=6}^{10} {}_{10}C_x \cdot 0.5^x \cdot (1-0.5)^{10-x}$$

$$= \frac{210+120+45+10+1}{2^{10}} \approx 0.377 > 0.5$$

同一性と可換性

- 同一性(identification)
  - すべてのデータx ($\forall x \in \Omega$)に対して、パラメータθ_a, θ_bが与えられた条件の下で$P(x|\theta_a) = P(x|\theta_b)$が成立するとき、パラメータは同一$\theta_a \equiv \theta_b$でなくてはならない。

- 可換性(exchangeability)
  - n回のベルヌーイ試行において、成功する回数sが同じであれば、成功・失敗が生じる順序がどうであれ確率は同じである

デ・フィネッティの定理

- 事前分布がどのような確率密度分布であっても、どの順番で事前分布が出現しても、十分長い期間ベイズの定理を適用すると、同じ事後分布に収束する

- 可換性の性質が示されたことにより尤度と事前情報の存在が示された

継起の規則（ラプラス）

- 何も情報がなければ、生まれてくる子供の性別が男である確率を$1/2$としてよいだろうか？
- n人の子供を産むとして、男の子か女の子が生まれる（死産は考えない）
- 生まれてくる子供の性別比率は事前にわからない
- 生まれてくる子供の性別が男である確率は

$$1/(n+1)$$

- 数学的帰納法により証明可能

継起の規則（ラプラス）

- $n=1$のとき、明らか
- $n>1$のとき、$n+1$人中m人（$m=0,\ldots,n+1$）が男の子である確率は、最初のn人でm人の男の子が生まれ、$n+1$人目は女の子が生まれることだから、$[(n-m)+1]/[(n+1)(n+2)]$となる
- それまで$m-1$人の男の子が生まれ、次に生まれるのが女の子である確率は、 $[(m-1)+1]/[(n+1)(n+2)]$となる
- この二つを足すと、$1/(n+2)$となる　■

継起の規則とベイズの定理

- ところで、男の子あるいは女の子に会う確率は、最初は五分五分（確率1/2）と考えるのが一般的（経験・経験による事前確率）
  - 主観的に事前確率は均一と考える

- 事前分布は均一？ どう設定すればいい？
  →　継起の規則とベイズの定理を使って検証
  - ずっと男の子ばかり生まれる状況を想定

継起の規則とベイズの定理

- 最初に生まれるn人が男の子である：事象A
- 次に産まれるm人が男の子である：事象B
- ベイズの定理より、

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} = \frac{1/(n+m+1)}{1/(n+1)} = \frac{n+1}{n+m+1}$$

事前（主観）確率は意味がある

- $n \to \infty$、$m \to \infty$なら、$P(B|A) \to 0$
- ずっと男の子が生まれたから、次は女の子が生まれるだろう、という推測は成立しない
- 事象Aに対するある程度の疑いは克服できない

- 事前確率の大小が問題ではない
- 事前確率が主観によることを認める意味で重要

理由不十分の原則
(Principle of insufficient reason)

- ある事前確率を与えたとき、事前確率がそれ以外考えられない場合、当面それを事前確率と認めてよいこと

- 例：
  - 生まれてくる子供が男の子である確率$= 1/2$
  - 薬が特効薬である場合に、薬を投与して病気が治癒する確率$= 1$

薬の効用の事後分布

- 5人の治験(＝ベルヌーイ試行)結果から、事後分布は次式で得られる

$$
\begin{aligned}
&p(y_1, y_2, y_3, y_4, y_5) \\
&\propto \ell(y_1|\theta)\cdot\ell(y_2|\theta)\cdot\ell(y_3|\theta)\cdot\ell(y_4|\theta)\cdot\ell(y_5|\theta) \\
&\propto \int_0^1 \theta^5(1-\theta)^{5-5}p(\theta)d\theta
\end{aligned}
$$

- ベルヌーイ試行の事後分布は次式を計算することで得られる

$$p(y_1,\cdots,y_n) \propto \int_0^1 \theta^n(1-\theta)^{n-s}p(\theta)d\theta$$

薬の効用（１）

- では、5人中4人が治癒した場合はどうか？
- 薬を投与して治癒する確率：$\ell(y|\theta)=\theta$
- 薬を投与して治癒しない確率：$\ell(\bar{y}|\theta)=1-\theta$

- 一人目：治癒した（理由不十分の原則）
  - 事前分布：$p(\theta)=1$
  - 事後分布：$\ell(y_1|\theta)\cdot p(\theta)\propto\theta\times1=\theta$
- 二人目：治癒した（薬が有効だった）
  - 事前分布： $p(\theta|y_1)=2\theta \quad \because \int_0^1 2\theta d\theta=[2\cdot(1/2)\theta^2]_0^1=1$
  - 事後分布： $\ell(y_2|\theta)\cdot p(\theta)=\theta\cdot2\theta=2\theta^2$

薬の効用（２）

- 三人目：治癒した
  - 事前分布： $p(\theta|y_2) = 3\theta^2 \quad \because \int_0^1 3\theta^2 d\theta = [3 \cdot (1/3)\theta^3]_0^1 = 1$
  - 事後分布： $\ell(y_3|\theta) \cdot p(\theta) = \theta \cdot 3\theta^2 = 3\theta^3$
- 四人目：治癒した
  - 事前分布： $p(\theta|y_3) = 4\theta^3 \quad \because \int_0^1 4\theta^3 d\theta = [4 \cdot (1/4)\theta^4]_0^1 = 1$
  - 事後分布： $\ell(y_4|\theta) \cdot p(\theta) = \theta \cdot 4\theta^3 = 4\theta^4$
- 五人目：治癒しなかった
  - 事前分布： $p(\theta|y_4) = 5\theta^4 \quad \because \int_0^1 5\theta^4 d\theta = [5 \cdot (1/5)\theta^5]_0^1 = 1$
  - 事後分布： $\ell(y_5|\theta) \cdot p(\theta) = 30(1-\theta) \cdot \theta^4$

薬の効用（３）

- この薬は「効く$\theta > 0.5$」といえるか

$$\int_{0.5}^{1} 30(1-\theta) \cdot \theta^4 d\theta = [30 \cdot \{(1/5)\theta^5 - (1/6)\theta^6\}]_{0.5}^1$$

$$= [6\theta^5 - 5\theta^6]_{0.5}^1 = 1 - \frac{7}{64} = \frac{57}{64} \approx 0.89$$

- この結果からは、新薬が有効といえそう

薬の効用(別の視点から)

- 新型インフルエンザの新薬：$1/2$の確率で治癒
- この新薬が特効薬（＝全員が治る）である可能性は？
- 古典統計的発想：p値

$$\left(\frac{1}{2}\right)^5 = \frac{1}{32} = 0.03125$$

- ベイズ的発想：特効薬なら全員が治癒する確率＝ 1
  （薬の事前効用は1）

$$1 - \left(\frac{1}{2}\right)^5 \approx 0.9697$$