数理統計 補助資料
～事前分布と事後分布～

2023年度2学期:　　月曜1限, 水曜3限

担当教員:　　　　石垣　司

ベイズ統計モデルのパラメータの推定

尤度の観点からみたベイズの定理の意味(前回授業の復習)

– **事後分布 $p(\boldsymbol{\theta}|\boldsymbol{y})$ はデータ観測後のパラメータの分布**

$$p(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{\prod_{i=1}^{N} p(y_i;\boldsymbol{\theta})\,p(\boldsymbol{\theta})}{\int p(\boldsymbol{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta}} \propto \prod_{i=1}^{N} p(y_i;\boldsymbol{\theta})\,p(\boldsymbol{\theta})$$

– **ベイズの定理の構造**

事後分布 ⇐ データ(尤度関数)× 事前分布

今回の授業のテーマ:
事前分布 $p(\boldsymbol{\theta})$ と事後分布 $p(\boldsymbol{\theta}|\boldsymbol{y})$ って何なの?

#メモ　これ以降、話が詳細に入っていくが目的を見失わないように。統計モデル(関数)の形を決めるパラメータの値を知りたくて、そのパラメータの値の推定を行うのが目的。

事前分布 $p(\boldsymbol{\theta})$

パラメータに関して事前に持っている情報

– **事前に持っている分析者の主観, 信念, 確信度などを確率分布によって表現**

– **ベイズ統計学特有の情報の表現方法**

– **不確実性を表現する確率の種類(客観確率と主観確率)**

1. **ワクチンとインフル発症率: 実験やデータに基づいた確率**
2. **低気圧と頭痛発生率: 過去の個人の経験から得た感触**
3. **告白と結果: 未経験の事象に対する信念**

頻度主義的統計学での確率は統計的確率 $P(A) = \lim_{n\to\infty} \frac{r}{n}$ で定義。2と3は確率として取り扱うことができない & 取り扱わない
ベイズ統計学では個人の確信度も確率(主観確率)として取り扱う

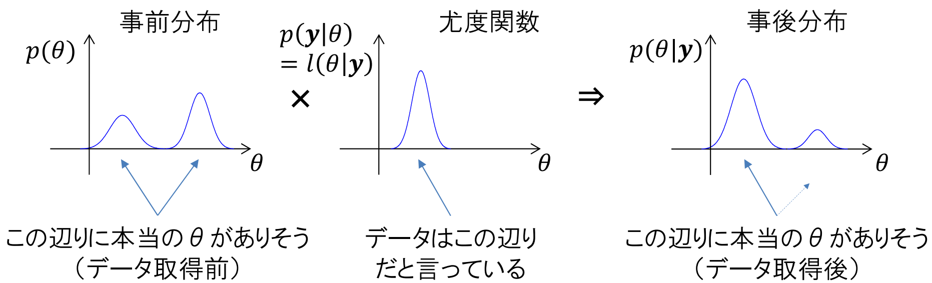
#メモ　本授業では詳しくは触れないが、客観確率と主観確率に関する確率に対しての哲学と客観性の重視度合(科学的立場・姿勢)が頻度主義的統計学とベイズ統計学では相容れなかったために論争が続いていた。

事後分布 $p(\boldsymbol{\theta}|\boldsymbol{y})$ のイメージ

データによって更新されたパラメータの情報

– **事前分布: データが得られる前のパラメータに関する情報を確率分布によって表現**

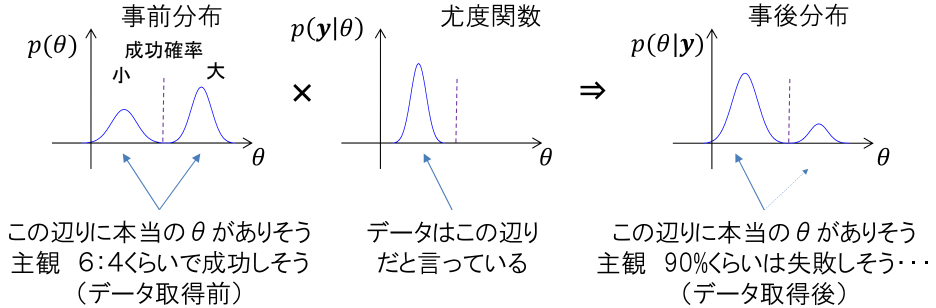
– **尤度: データ自体が持つパラメータに関する情報**



主観確率と事後分布の例 #1

未経験の事象の確率的表現～告白の成功確率の例

- **θ は相手からの好意度に関連するパラメータ**

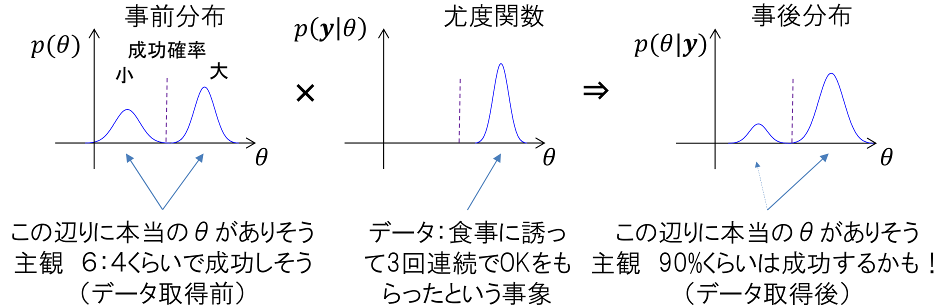


- **事前分布：告白の成功確率について持っている主観（確信度）**
- **尤度関数：3回食事に誘ってすべて断られたという事象（データ）**
- **事後分布：成功の可能性も残っているが失敗しそう，という主観（確信度）**

主観確率と事後分布の例 #2

未経験の事象の確率的表現～告白の成功確率の例

- **θ は相手からの好意度に関連するパラメータ**



ベイズ統計学の有用性の1つ

- **人間の思考や意思決定プロセスとベイズの定理（事後分布の算出）がアナロジーとなり，現実の問題に適用しやすい**

事後分布の例～正規分布 #1

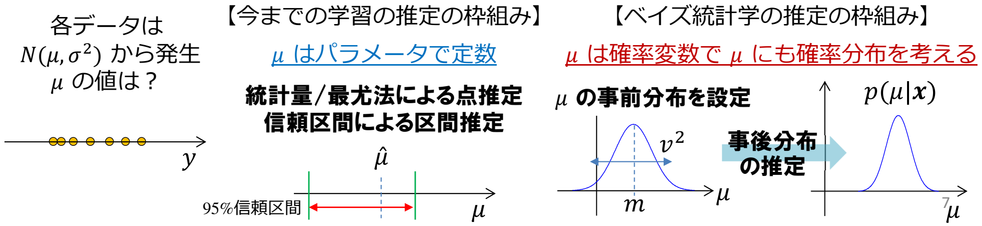
正規分布 $N(\mu,\sigma^2)$ の分散 σ^2 は既知として，データ $\{y_i\}$ が観測されたときの平均 μ の事後分布を導出

条件1：データ y_i は正規分布から発生

- **データ** $y_i \sim i.i.d.\, N(\mu,\sigma^2), (i=1,\ldots,N)$

条件2：事前分布は正規分布と設定

- **平均 μ の事前分布** $p(\mu)=N(\mu;m,v^2)$ **を設定**



事後分布の例～正規分布 #2

$N(\mu;m,v^2)$：μ に関する平均 m，分散 v^2 の正規分布の確率密度関数

μ の事後分布の導出

- **平均 μ の事前分布：正規分布を設定**

$$p(\mu)=N(\mu;m,v^2)=\frac{1}{\sqrt{2\pi v^2}}\exp\left\{-\frac{1}{2v^2}(\mu-m)^2\right\}$$

- **尤度関数：**

$$p(\boldsymbol{y}|\mu)=l(\mu|\boldsymbol{y})=\prod_{i=1}^{N}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{1}{2\sigma^2}(y_i-\mu)^2\right\}$$

- **平均 μ の事後分布**

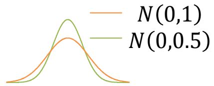
$$p(\mu|\boldsymbol{y})=\frac{1}{\sqrt{2\pi\dfrac{v^2\sigma^2}{Nv^2+\sigma^2}}}\exp\left\{-\frac{1}{2\dfrac{v^2\sigma^2}{Nv^2+\sigma^2}}\left(\mu-\frac{N\bar{y}v^2+m\sigma^2}{Nv^2+\sigma^2}\right)^2\right\}$$

$$=N\left(\mu;\frac{N\bar{y}v^2+m\sigma^2}{Nv^2+\sigma^2},\frac{v^2\sigma^2}{Nv^2+\sigma^2}\right)$$ check!

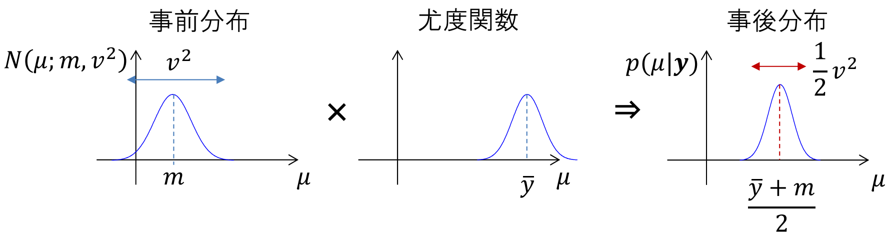
事後分布の例～正規分布 #3

例：$\sigma^2 = v^2 = 1$ かつ $N = 1$ のときの事後分布

$$p(\mu|\boldsymbol{y}) = N\left(\mu; \frac{\bar{y}+m}{2}, \frac{1}{2}\right)$$



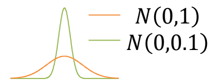
- **事後分布の平均は事前分布と尤度の中間点**
- **事後分布の分散は事前分布の半分になる（推定の精度が向上している）**



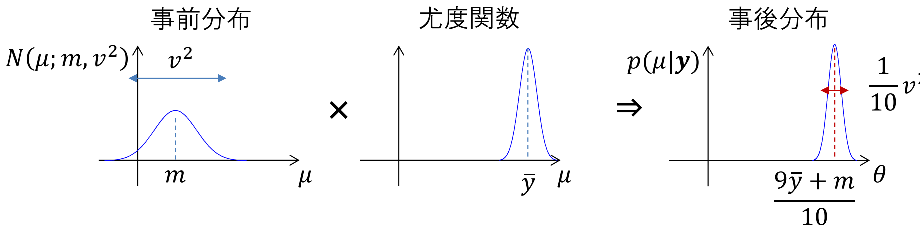
事後分布の例～正規分布 #4

例：$\sigma^2 = v^2 = 1$ かつ $N = 9$ のときの事後分布

$$p(\mu|\boldsymbol{y}) = N\left(\mu; \frac{9\bar{y}+m}{10}, \frac{1}{10}\right)$$



- **事後分布の平均はサンプルサイズ N に比例して尤度の情報が支配的になっていく**
- **事後分布の分散はサンプルサイズ N に反比例して小さくなる（推定の精度が向上する）**



事後分布の例～線形回帰モデル #1

例：単回帰モデル

$$y = a + bx + e, e \sim N(0, \sigma^2)$$

- **分散 σ^2 と切片 a は既知で，傾き b のみの推定を考える**

- **傾き b の事前分布（正規分布を採用）**

$$p(b) = N(b; m, v^2) = \frac{1}{\sqrt{2\pi v^2}} \exp\left\{-\frac{1}{2v^2}(b-m)^2\right\}$$

- **尤度関数**

  目的変数：$\boldsymbol{y} = [y_1, \ldots, y_N]^T$ (i. i. d.). **説明変数：**$\boldsymbol{x} = [x_1, \ldots, x_N]^T$

  データ：$D = \{\boldsymbol{y}, \boldsymbol{x}\}$

$$p(y_i|x_i; \boldsymbol{\theta}) = N(a + bx_i, \sigma^2)$$

$$p(D|b) = l(b|D) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}\{y_i - (a + bx_i)\}^2\right]$$

事後分布の例～線形回帰モデル #2

例：単回帰モデル

$$y = a + bx + e, e \sim N(0, \sigma^2)$$

- **分散 σ^2 と切片 a は既知で，傾き b のみの推定を考える**

- **傾き b の事後分布**

check!

$$p(b|D) = N\left(b; \frac{(\boldsymbol{y}-a)^T \boldsymbol{x} v^2 + m\sigma^2}{\boldsymbol{x}^T \boldsymbol{x} v^2 + \sigma^2}, \frac{\sigma^2 v^2}{\boldsymbol{x}^T \boldsymbol{x} v^2 + \sigma^2}\right)$$

#メモ　本授業では、正規分布や線形回帰モデルの事後分布の推定についてこれ以上の紹介しないが、分散 σ^2 にも事前分布を設定した場合の事後分布、重回帰モデル回帰係数の事後分布などの導出などが代表的な学習の道筋である。

演習問題

あなたはある製造機器メーカーのマネージャーとする。自社の新製品を市場に投入する前には, その新製品の販売台数は1か月につき平均が10, 分散が1の正規分布にしたがうという確信があった。

しかし, 市場投入後1か月目の販売台数は16台, 市場投入後2か月目の販売台数は19台であった。市場の構造や製品の認知度などすべての内部環境と外部環境に変化がない, かつ, 各月の販売台数は分散が1の正規分布に従うと仮定したとき,ベイズ統計学に立脚して販売台数の事後分布を求めて, 次月(市場投入後3か月目)の販売台数を予測しなさい