

情報数学

中山クラス 第13週

<今日の内容>

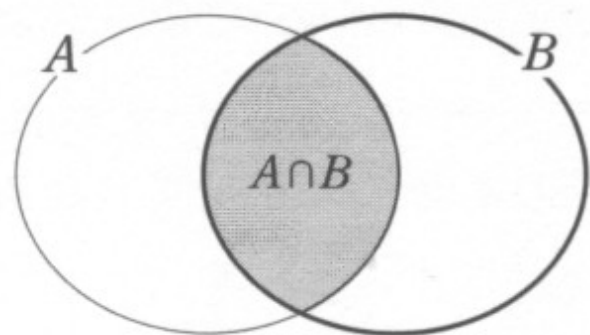
- ◇小テストの解説
- ◇第3章 ベイズ統計学の基本
 1. ベイズ統計はシンプルな最強ツール
 2. ベイズ統計の基本公式
 3. コインの問題を考える
- ◇演習問題

3章 ベイズ統計学の基本

- 統計モデルに含まれる母数(パラメータ)をベイズの定理に取り込む.
- ベイズ統計の基本を調べる.
- 母数が確率変数として扱われる.

1 ベイズ統計はシンプルな最強ツール p.76

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$



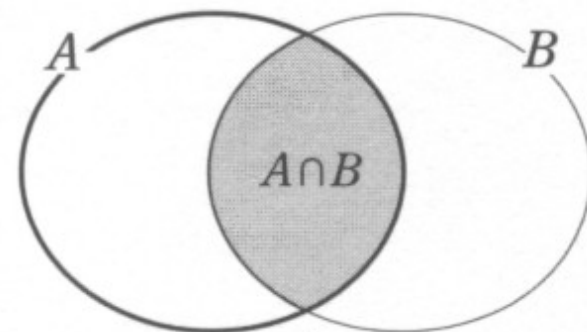
$P(A|B)$
(Bが起こったときのAの確率)

ベイズの公式

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

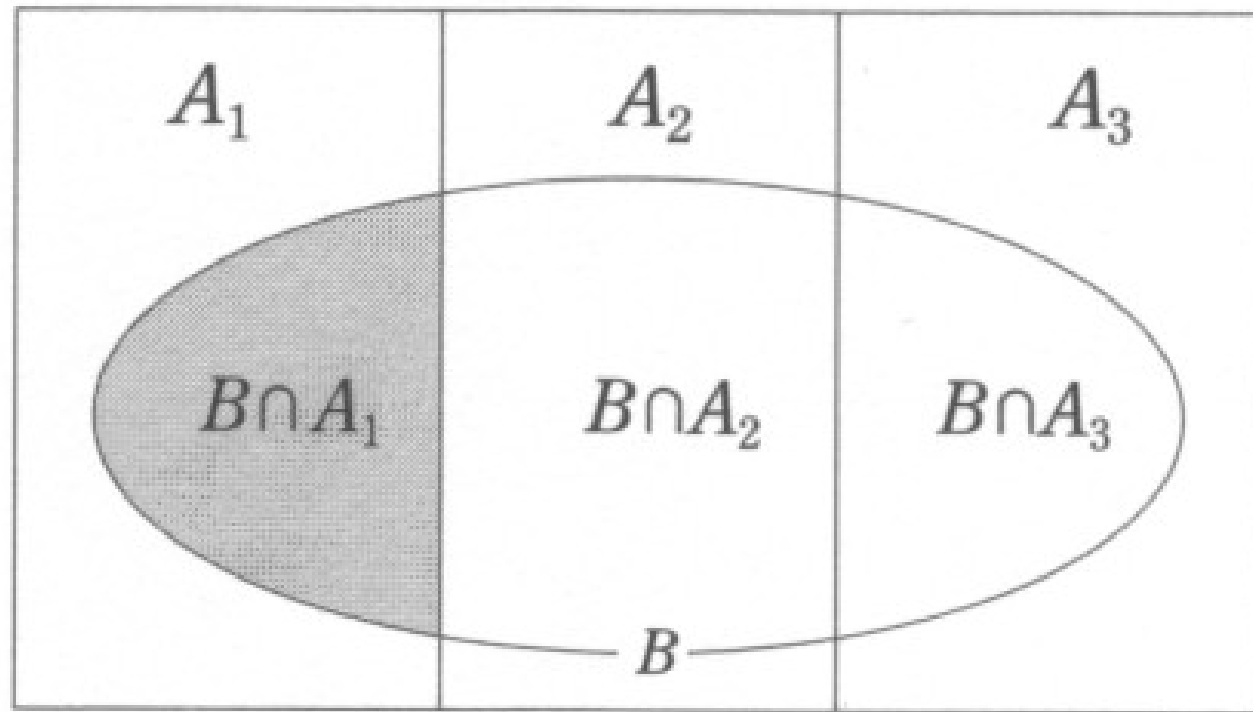


変換



$P(B|A)$
(Aが起こったときのBの確率)

$$P(A_1|B) = \frac{P(B|A_1)P(A_1)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3)}$$



■ ベイズの定理の変形

ベイズの定理において

A, B 確率現象の事象であればなんでも良い.

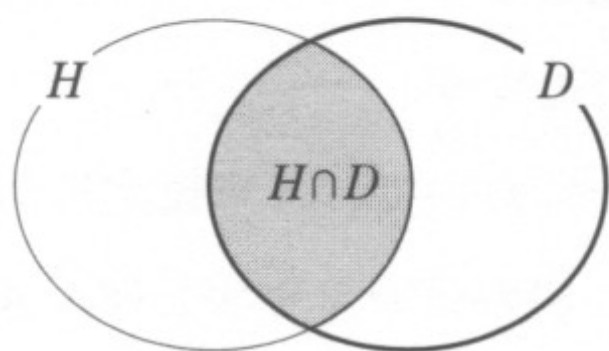


A: 仮定に関する事象 → 原因 → H (Hypothesis)

B: その結果に関する事象 → データ → D (Data)

$$P(H|D) = \frac{P(D|H)P(H)}{P(D)}$$

データDが得られた後に、その原因となる仮定Hが成立していた確率(原因の確率)を求める公式



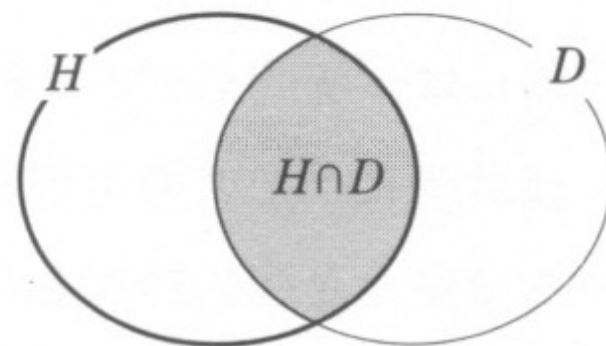
データDが得られたときに
原因がHである確率

ベイズの公式

$$P(H|D) = \frac{P(D|H)P(H)}{P(D)}$$



変換

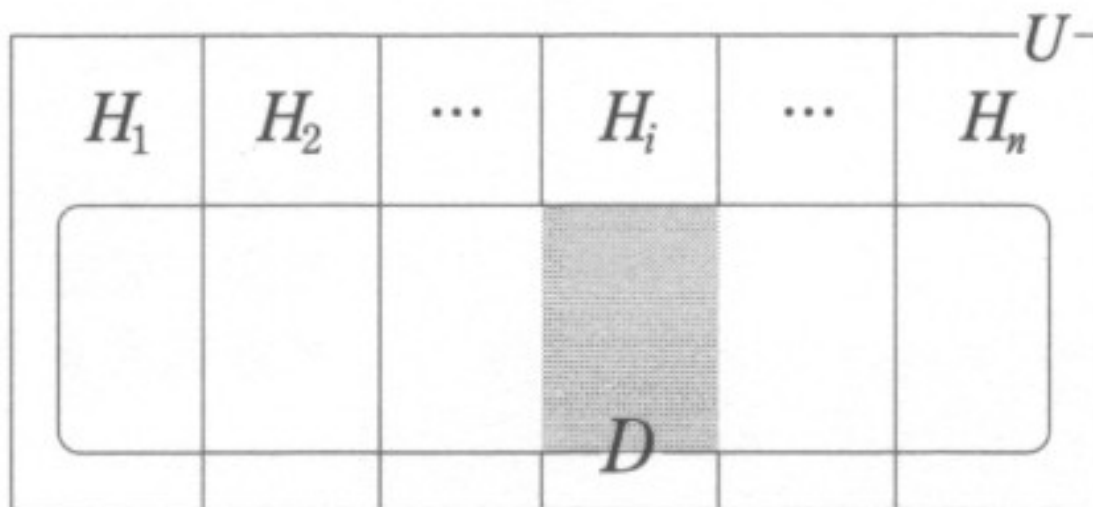


原因がHのときに
データDが得られる確率

仮定Hがいろいろある場合のベイズの公式

データDが得られた後に、その原因となる仮定が H_i である確率を表す。

$$P(H_i|D) = \frac{P(D|H_i)P(H_i)}{P(D|H_1)P(H_1) + P(D|H_2)P(H_2) + \dots + P(D|H_n)P(H_n)}$$



データ D は仮定 H_1 、 H_2 、 $\dots$ 、 H_n のどれかが原因で得られるとする。

■例題(壺の問題) p.79

(問題)

1個の壺がある. 壺の中には白と赤の3個の玉が入っている. そこから玉1個を取り出したとき, それが赤玉であった(結果). 壺の中に入っている赤玉の個数(仮定／原因)の確率を求めよ.

<解答例>

仮定(原因): 壺の中の赤玉の個数 = 1, 2, 3個

壺1 [○ ○ ●], 壺2 [○ ● ●], 壺3 [● ● ●]

H_k : 壺 k から玉1個取り出す. $k = 1, 2, 3$

結果: 取り出した玉が赤玉である.

D : 壺から玉1個を取り出したとき, それが赤玉である.

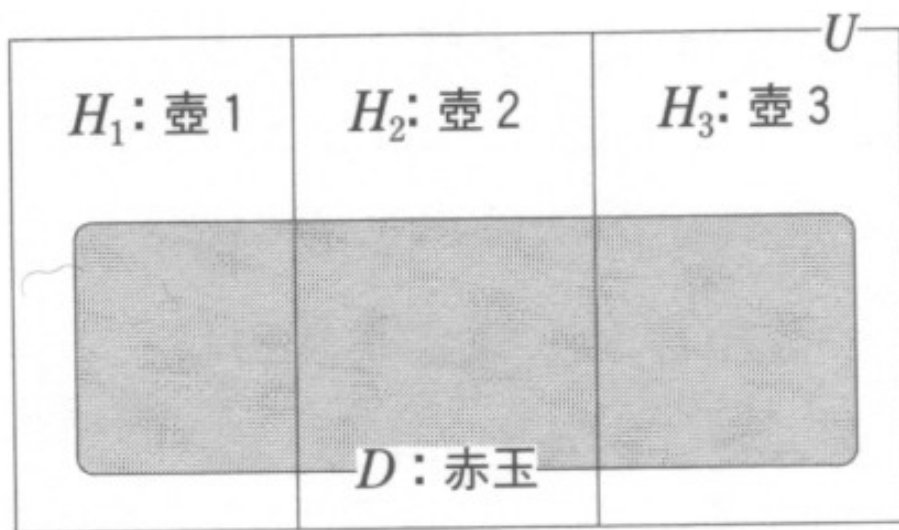
目標

赤玉が得られたとき, それが壺 i から取り出された確率を全ての $i = 1, 2, 3$ について求める.



データ D が得られたとき, その仮定が H_i である確率 $P(H_i|D)$ を全ての $i = 1, 2, 3$ について求める.

$$P(H_i|D) = \frac{P(D|H_i)P(H_i)}{P(D|H_1)P(H_1) + P(D|H_2)P(H_2) + P(D|H_3)P(H_3)}$$



$P(D|H_i)$: 3個中, 赤玉が i 個入った壺から赤玉を1個取り出す確率.

$$P(D|H_1) = 1/3, P(D|H_2) = 2/3, P(D|H_3) = 3/3$$

$P(H_i)$: 壺 i が選ばれる確率(問題では与えられていない)
→「理由不十分の原則」に基づき等確率とする.

$$P(H_1) = P(H_2) = P(H_3) = 1/3$$

➤ 壺の選択に関して条件が与えられていれば, それを考慮することが出来る→ベイズ統計のメリット

上記の確率を用いて目的の確率分布が求まる(pp.80-81).

赤玉の個数	1個	2個	3個
-------	----	----	----

確率	$P(H_1 D) = 1/6$	$P(H_2 D) = 1/3$	$P(H_3 D) = 1/2$
----	------------------	------------------	------------------

確率の合計: $1/6 + 1/3 + 1/2 = 1$

■ 尤度, 事前確率, 事後確率 p.81

$$P(H_i|D) = \frac{P(D|H_i)P(H_i)}{P(D|H_1)P(H_1) + P(D|H_2)P(H_2) + P(D|H_3)P(H_3)}$$

$P(H_i)$: 玉を取り出す前に壺が選ばれる確率 → 事前確率

$P(H_i|D)$: 赤玉が取り出された後に, それが壺*i*から取り出された確率 → 事後確率

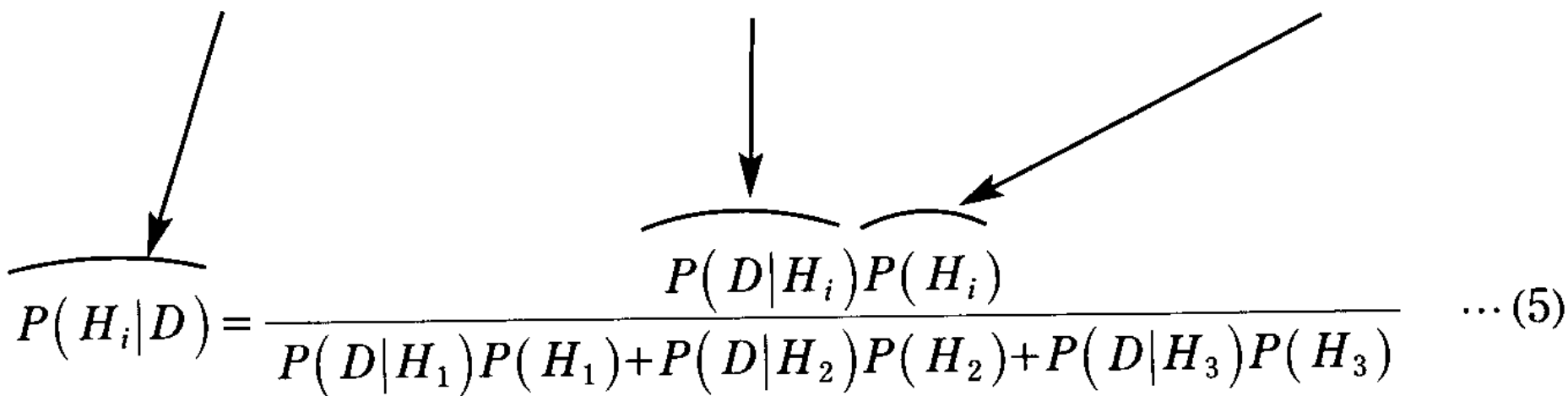
$P(D|H_i)$: 壺*i*が選択されたとき, そこから赤玉が取り出される確率 → 仮定*H_i*の尤度

p.82

赤玉が取り出された後に、それが壺 i のものである確率 (事後確率)

壺 i が選択されたときに、そこから赤玉が取り出される確率 (尤度)

玉を取り出す前に壺 i が選択される確率 (事前確率)


$$P(H_i|D) = \frac{P(D|H_i)P(H_i)}{P(D|H_1)P(H_1) + P(D|H_2)P(H_2) + P(D|H_3)P(H_3)} \dots (5)$$

データ D が得られたときに、その原因が H である確率（事後確率）

原因が H であるときに、データ D が得られる確率（尤度）

原因 H が発生する確率（事前確率）

ベイズの公式

$$P(H|D) = \frac{P(D|H)P(H)}{P(D)} \dots (3)$$

2 ベイズ統計の基本公式 p.83

■母数(パラメータ)の導入

統計学では母数(パラメータ)を推定することが重要
母数の例

正規分布では「平均値 μ 」, 「分散 σ^2 (標準偏差 σ)」

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

★ベイズの定理で母数を推定

仮定 $H \rightarrow$ 「母数が θ の値をとる」

文章としての仮定 H を数値としての θ と解釈する.

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}$$

■ 母数が連続変数の場合のベイズの定理

p.84

確率変数が離散的な値を取る場合

$P(\theta|D), P(D|\theta), P(\theta)$ は確率そのものを表す.

確率変数が連続的な値を取る場合

θ がある値を取る確率は零である.

確率分布を表すもの＝確率密度関数

確率＝確率密度関数の積分

ベイズの定理においては、確率を確率分布と読み替える.

事前確率 $P(\theta)$ $\rightarrow$ 事前(確率)分布 $\pi(\theta)$

尤度 $P(D|\theta)$ $\rightarrow$ 尤度 $f(D|\theta)$

事後確率 $P(\theta|D)$ $\rightarrow$ 事後(確率)分布 $\pi(\theta|D)$

ベイズの定理

- θ が離散的な値を取る場合

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}$$

- θ が連続的な値を取る場合

$$\pi(\theta|D) = \frac{f(D|\theta)\pi(\theta)}{P(D)}$$

データ D が得られたときの、母数 θ の確率密度関数 (事後分布)

原因が H であるときにデータ D が得られる確率 (尤度)

母数 θ の確率密度関数 (事前分布)

ベイズの定理

$$\pi(\theta|D) = \frac{f(D|\theta)\pi(\theta)}{P(D)} \dots (3)$$

■さらにコンパクトに p.85

ベイズの定理において

$P(D)$: データ D が得られる確率

ベイズ統計ではデータ D が得られた後のことを考える.
母数の推定には直接影響しない.

$P(D)$ を定数として扱う.

$$P(\theta|D) = kP(D|\theta)P(\theta)$$

$$\pi(\theta|D) = kf(D|\theta)\pi(\theta)$$

k は次の規格化条件より求まる.

$$\sum_{\theta=1}^L P(\theta|D) = 1, \quad \int_{-\infty}^{\infty} \pi(\theta|D)d\theta = 1$$

ベイズ統計の基本公式 p.86

比例定数 k を明示するのが面倒な場合もある.
→比例関係で表現する.

ベイズ統計の基本公式

事後分布は尤度と事前分布の積に比例する.

事後分布 $\pi(\theta|D) \propto$ 尤度 $f(D|\theta) \times$ 事前分布 $\pi(\theta)$

★ θ が複数の母数を表す場合にも基本公式を適用できる.

例題(製造ラインにおける平均値の分布)p.87

(問題)

- 菓子Aの製造ラインから作られる製品の重さの平均値 μ を調べる.
- 3個の製品を取りだしたところ, 99g, 100g, 101gであった.
- これまでの検査によって, このラインから製造される製品の重さの分散は3であることが分かっている.
- また, 去年の経験から, 平均値 μ は平均値100, 分散1の正規分布に従っていると想像される.
- ◆ このとき, 菓子Aの重さの平均値 μ の事後分布を求めよ.

<解答例> pp.87-88

同じ製造ラインで作られる製品の重さは正規分布になる場合が多い. この問題でも正規分布で考える.

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

正規分布の母数: 平均値 μ (調べる対象), 分散 $\sigma^2 = 3\text{g}$

得られたデータ D : 3個の製品の重さ = 99g, 100g, 101g

これらのデータは正規分布(平均値 μ , 分散 $\sigma^2 = 3$)に従う母集団から取り出されたものである.

尤度: $f(D|\mu) =$

$$\frac{1}{\sqrt{2\pi \times 3}} e^{-\frac{(99 - \mu)^2}{2 \times 3}} \frac{1}{\sqrt{2\pi \times 3}} e^{-\frac{(100 - \mu)^2}{2 \times 3}} \frac{1}{\sqrt{2\pi \times 3}} e^{-\frac{(101 - \mu)^2}{2 \times 3}}$$

母数 μ の事前分布 $\pi(\mu)$

去年の経験から平均値100, 分散1の正規分布に従っていると仮定できる.

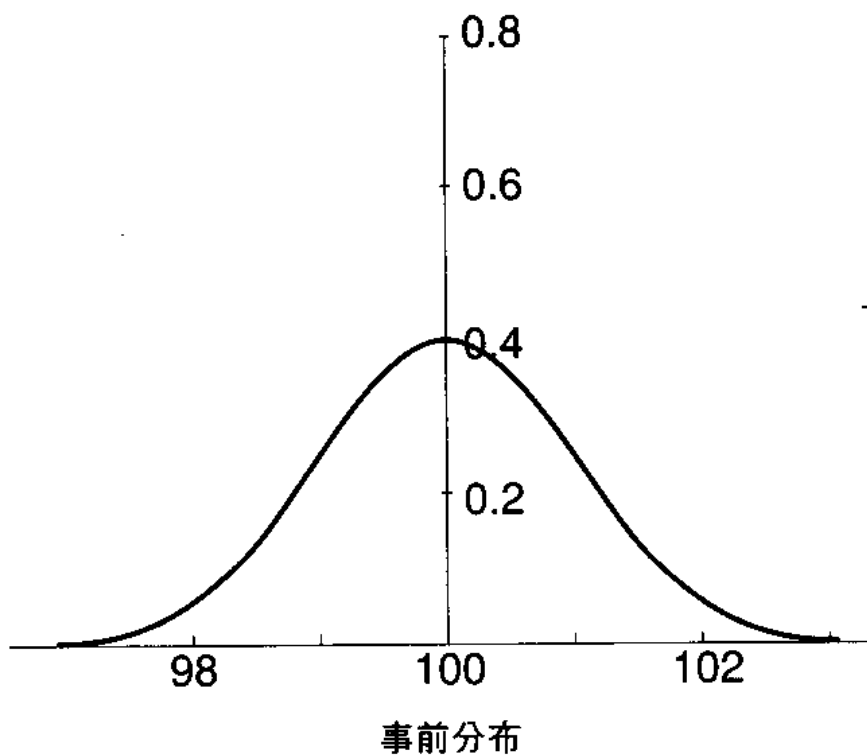
$$\pi(\mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(\mu-100)^2}{2}}$$

母数 μ の事後分布: $\pi(\mu|D)$

$$\begin{aligned}\pi(\mu|D) &\propto f(D|\mu) \times \pi(\mu) \\ &= \frac{1}{\sqrt{2\pi \times 3}} e^{-\frac{(99-\mu)^2}{2 \times 3}} \frac{1}{\sqrt{2\pi \times 3}} e^{-\frac{(100-\mu)^2}{2 \times 3}} \\ &\quad \times \frac{1}{\sqrt{2\pi \times 3}} e^{-\frac{(101-\mu)^2}{2 \times 3}} \frac{1}{\sqrt{2\pi}} e^{-\frac{(\mu-100)^2}{2}} \\ &\propto e^{-\frac{1}{2 \times \frac{1}{2}} (\mu-100)^2}\end{aligned}$$

平均値 μ の事前分布

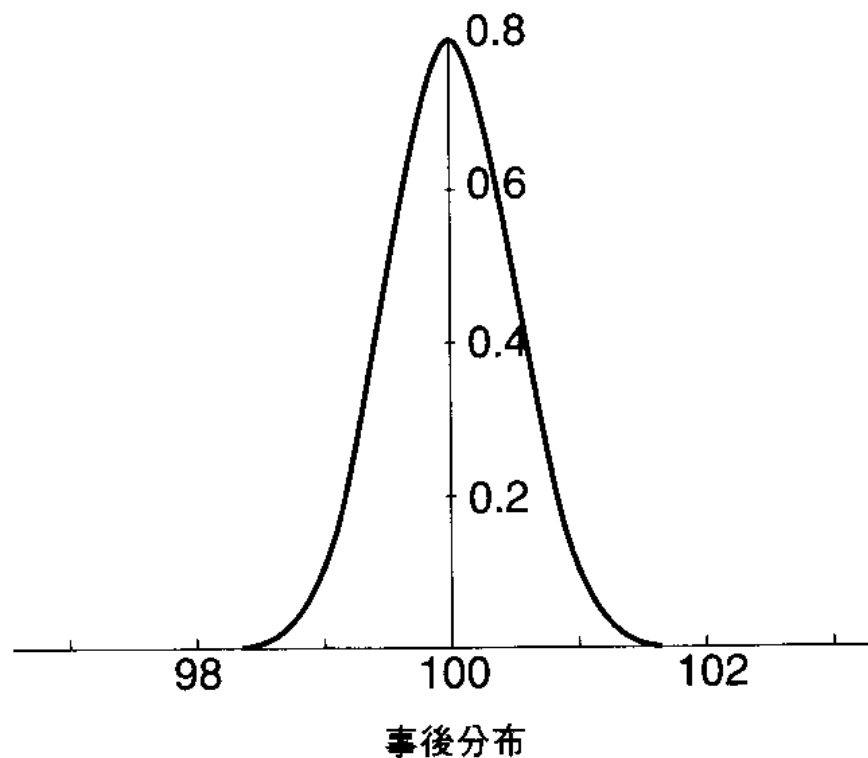
去年の経験から、平均値100, 分散1の正規分布に従うと想定された.



平均値 μ の事後分布

3個のデータを用いて、ベイズ統計解析することにより、平均値100, 分散1/2の正規分布に従うことが分かった.

$\mu = 100$ の精度が高くなった.



3 コインの問題 p.90

(問題) 表の出る確率が θ である1枚のコインがある. このコインを4回投げたとき, 次のようになった.

1回目: 表 → 2回目: 表 → 3回目: 裏 → 4回目: 裏

(コインを投げることによって得られる結果／データ)

このとき, 表の出る確率 θ の事後分布を求めよ.

<解答例>

対象となる母数: 表の出る確率 = θ

■ 尤度 $f(D|\theta)$: 「表の出る確率 = θ 」のもとで D (表／裏が出る) が起こる確率 (条件付き確率)

$$f(\text{表}|\theta) = \theta$$

$$f(\text{裏}|\theta) = 1 - \theta$$

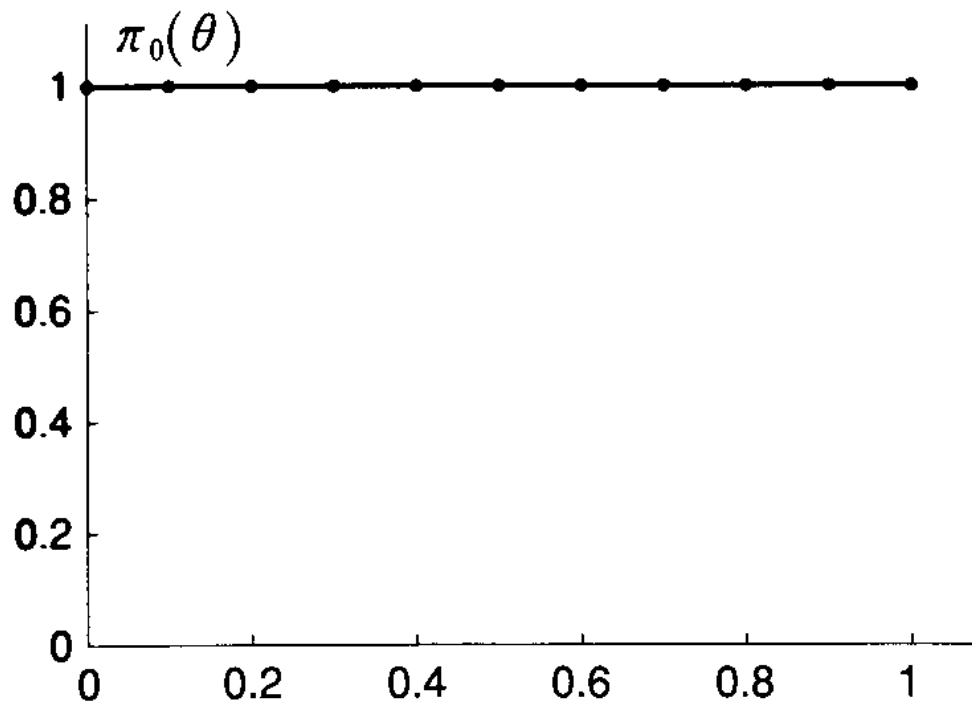
■事前分布: $\pi(\theta)$...確率密度関数 p.91

「表の出る確率」は $0 \leq \theta \leq 1$ の範囲で考えられる. この範囲で θ がどのように分布するかの事前情報はない.



「理由不十分の原則」に基づいて「一様分布」すると考える.

$$\pi_0(\theta) = 1, 0 \leq \theta \leq 1$$



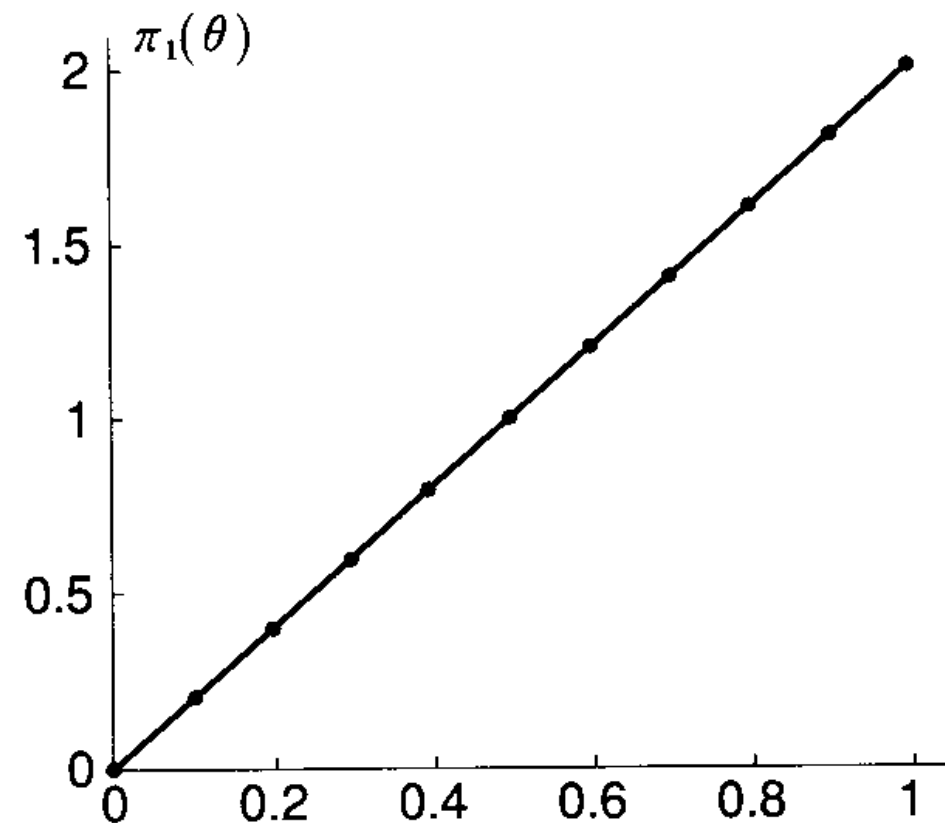
コインを投げていない
事前分布なので $\pi_0(\theta)$
とする.

■「1回目に表が出た」というデータを取り込む

D_1 : 1回目に表が出る.

コインを1回投げた後の θ の事後分布

$$\pi(\theta|D_1) \propto f(D_1|\theta) \times \pi_0(\theta) = \theta \times 1 = \theta$$



規格化条件(面積=1)より,
 $\pi_1(\theta) = \pi(\theta|D_1) = 2\theta$

コインを投げる前
情報がないので「一様分布」

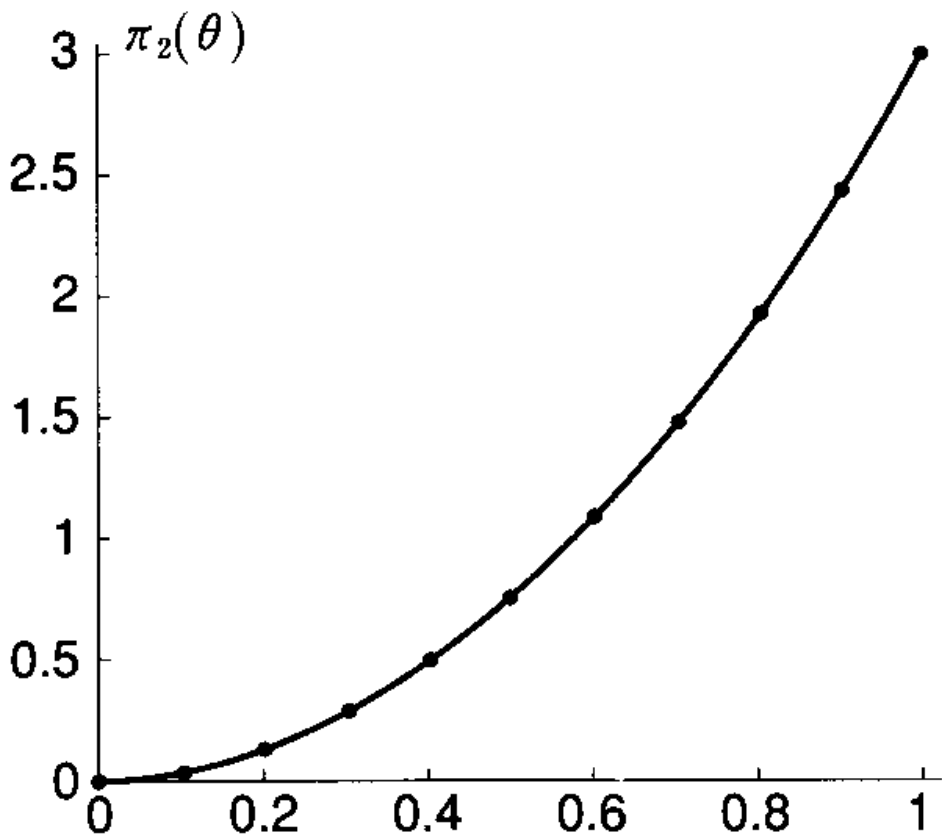
コインを1回投げた後
表が出たので,「表が出やすい分布」に更新された.

■「2回目に表が出た」というデータを取り込む

D_2 : 2回目に表が出る.

コインを2回投げた後の θ の事後分布

$$\pi(\theta|D_2) \propto f(D_2|\theta) \times \pi_1(\theta) = \theta \times 2\theta = 2\theta^2$$



規格化条件(面積=1)より,
 $\pi_2(\theta) = \pi(\theta|D_2) = 3\theta^2$

コインを1回投げた後
表が出たので、「表が出やすい分布 = 2θ 」に更新された.

コインを2回投げた後
表が出たので、「さらに表の出やすい分布 = $3\theta^2$ 」に更新された.

■「3回目に裏が出た」というデータを取り込む

D_3 : 3回目に裏が出る.

3回目に対する事前分布: $\pi_2(\theta) = 3\theta^2$

3回目に対する尤度: $f(D_3|\theta) = 1 - \theta$

コインを3回投げた後の θ の事後分布

$$\pi(\theta|D_3) \propto f(D_3|\theta) \times \pi_2(\theta) = (1 - \theta) \times 3\theta^2$$

規格化条件(面積=1)より,

$$\pi_3(\theta) = \pi(\theta|D_3) = 12(1 - \theta)\theta^2$$

■「4回目に裏が出た」というデータを取り込む

D_4 : 4回目に裏が出る.

4回目に対する事前分布: $\pi_3(\theta) = 12(1 - \theta)\theta^2$

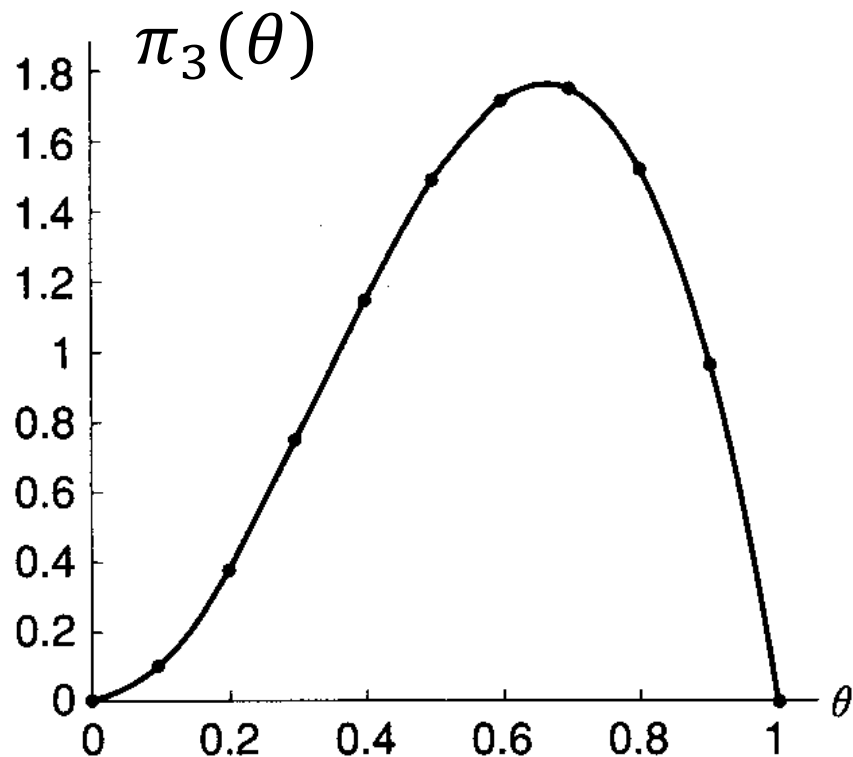
4回目に対する尤度: $f(D_4|\theta) = 1 - \theta$

コインを3回投げた後の θ の事後分布

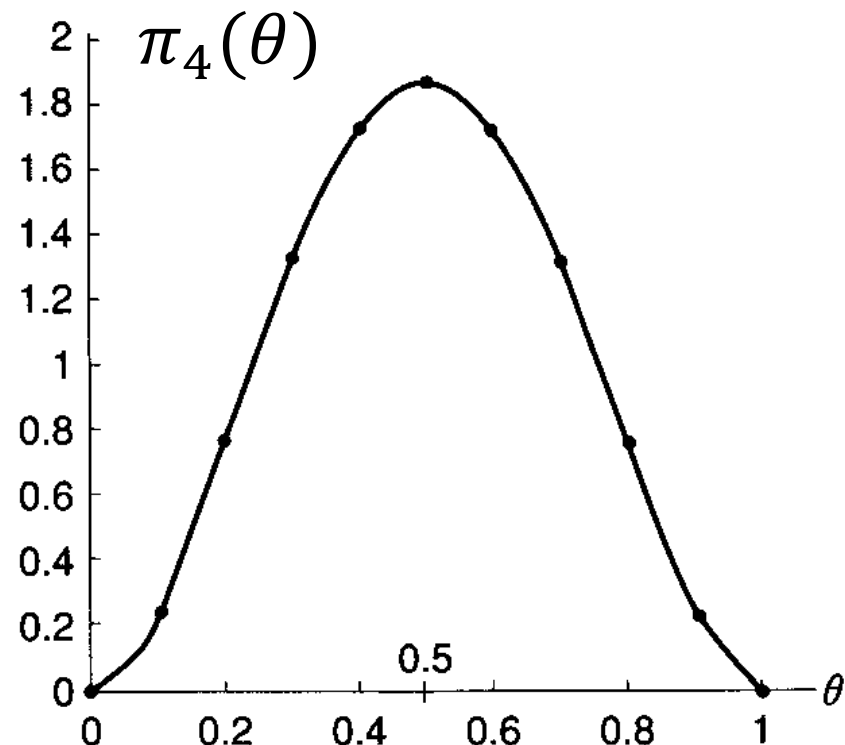
$$\pi(\theta|D_4) \propto f(D_4|\theta) \times \pi_3(\theta) = (1 - \theta) \times 12(1 - \theta)\theta^2$$

規格化条件(面積=1)より,

$$\pi_4(\theta) = \pi(\theta|D_4) = 30(1 - \theta)^2\theta^2$$



3回目では裏が出たため、 $\theta = 1$ (必ず表が出る) に対する事後分布は零になっている。しかし、まだ $\theta = 0.7$ 付近にピークがある。すなわち、表が出やすい分布になっている。



4回目で、再び裏が出たため、 $0.5 < \theta$ (表が出やすい) に対する事後分布が小さくなっている。表2回、裏2回となったため、 $\theta = 0.5$ 付近がピークとなり、表と裏の出やすさが均衡している。

ベイズ更新

データをつけ加えるごとに、母数 θ の確率分布が更新される・・・ベイズ統計の大きな特徴

新たなデータが加わったとき、はじめから統計解析をやり直す必要がなく、これまでの解析結果に新たなデータを追加する形で統計解析を更新できる。

■ ベイズ推定 p.94

ベイズ推定：事後分布から母数を推定する.

$\pi_4(\theta)$ を利用して, θ の平均値を推定する.

グラフが $\theta = 0.5$ を中心にして左右対称であることから,

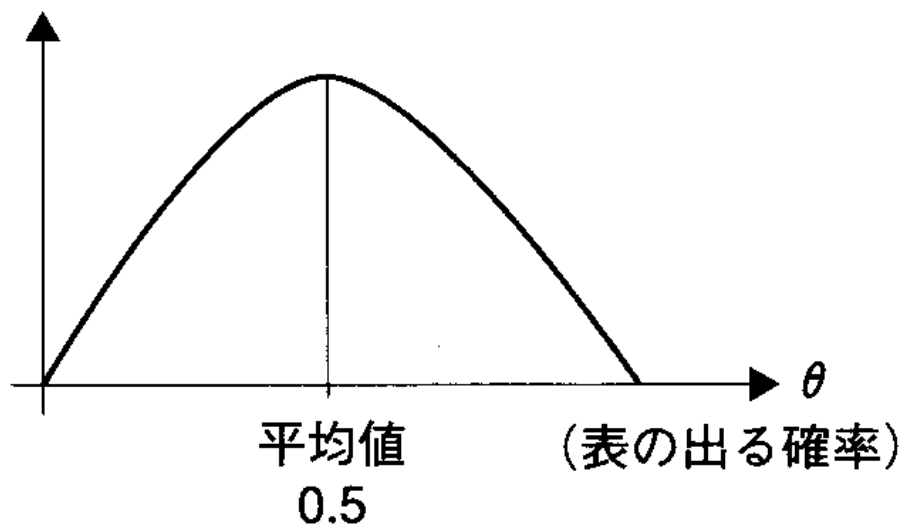
$$\theta = 0.5$$

が平均値である.

ベイズ統計の結果は
経験や直感と一致す
る！

表 表 裏 裏 →

事後分布



演習問題

1個の壺がある. 壺の中には白と赤の2個の玉が入っている. そこから玉1個を取り出したとき, それが赤玉であった(結果). 壺の中に入っている赤玉の個数(原因／仮定)に対する確率を求めよ.