情報数学

中山クラス
第１４週

＜今日の内容＞

◇第3章　ベイズ統計学の基本
　３．コインの問題を考える（第１３週のスライド）
　４．薬の効用問題

◇達成度確認試験の予想問題と解説

４　薬の効用問題　p.96

（例）新薬の効用
５人の治験者を抽出→４人に効き，１人には効かない
新薬の効き具合の分布を調べよ.

効き具合＝抽出された1人に新薬が効く確率
確率＝1（母集団全員に効く），＝0（誰にも効かない）

＜解説＞
ベイズ統計の母数＝１人の治験者に新薬が効く確率θ
■尤度：$f(D|\theta)$
「効く確率θ」のもとで，データD（５人のうち４人に効き，１人に効かない）の起こる確率（二項分布より求まる）

$$f(D|\theta) = {}_5C_4\theta^4(1-\theta)^1, 0 \le \theta \le 1$$

■事前分布：$\pi(\theta)$　p.96

治験するまでは効く確率θに関する情報はない.

「理由不十分の原則」より「全ての可能性は均等である」

$$\pi(\theta) = \text{定数}(k)$$

$0 \leq \theta \leq 1$であり, 確率の総和（面積）＝1であるから,

$$\text{事前分布}\quad \pi(\theta) = 1, 0 \leq \theta \leq 1$$

■事後分布：$\pi(\theta|D)$　p.97

ベイズ統計の基本公式（事後分布∝尤度×事前分布）より

$$\pi(\theta|D) \propto {}_5C_4\theta^4(1-\theta)^1 \times 1 \propto \theta^4(1-\theta)$$

$\pi(\theta|D)$の積分＝1より,

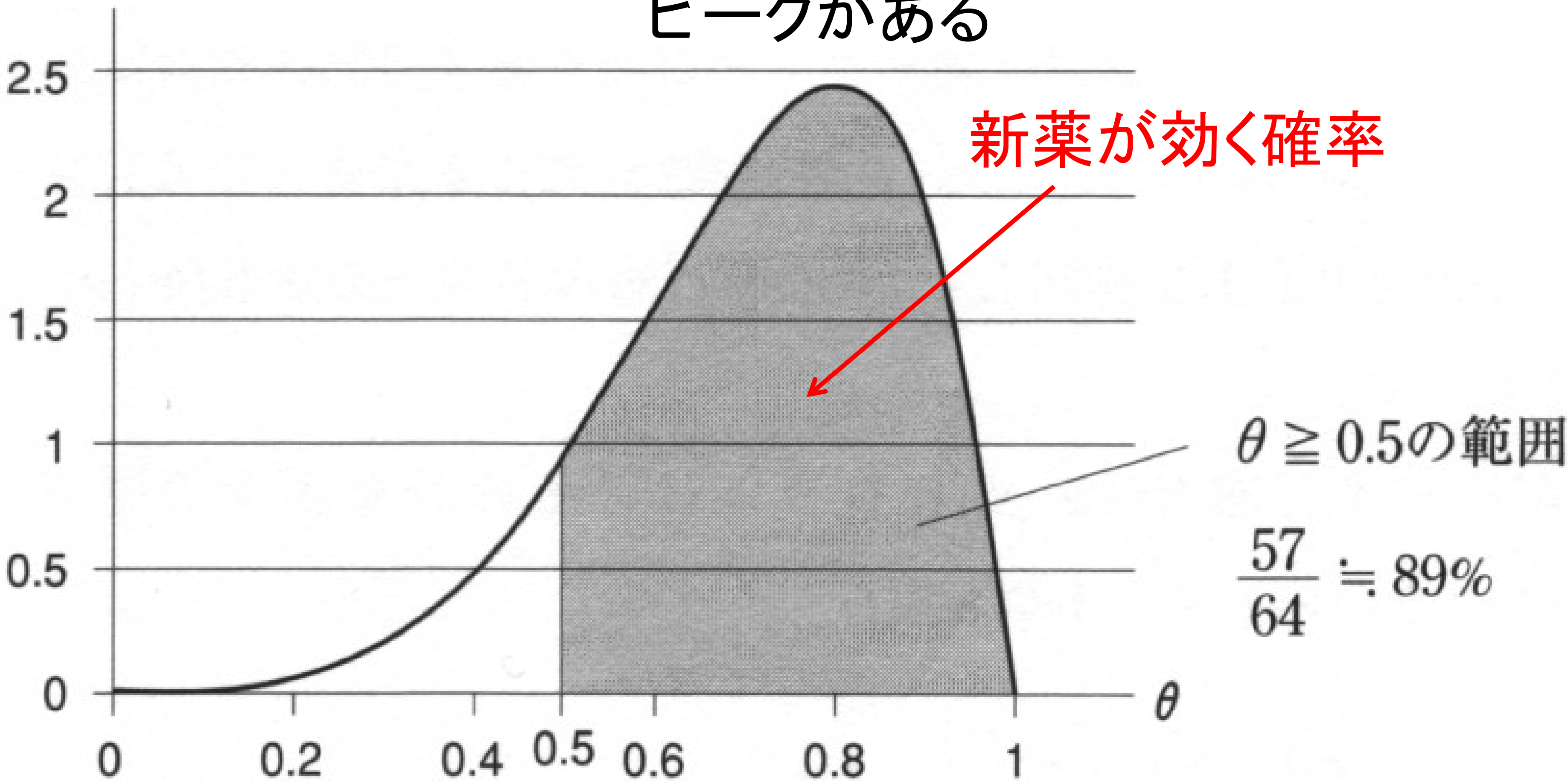
$$\pi(\theta|D) = 30\theta^4(1-\theta)$$

事後分布　$\pi(\theta|D) = 30\theta^4(1-\theta)$のグラフ

p.97

$\theta = 0.8$に
ピークがある

新薬が効く確率



■　分析してみよう　p.98
5人中4人に効果があった → 効く確率が高い（経験則）

事前分布　$\pi(\theta) = 1, 0 \leq \theta \leq 1$

データ（5人中4人に効く）

事後分布　$\pi(\theta|D) = 30\theta^4(1-\theta)$

確率の計算
薬が効く　→　$\theta \geq 0.5$
薬が効かない　→　$\theta < 0.5$

$$P(0.5 \leq \theta \leq 1) = \int_{0.5}^{1} \pi(\theta|D)d\theta \cong 0.89$$

平均値の計算

$$\mu = \int_{0}^{1} \theta\pi(\theta|D)d\theta = \frac{5}{7}$$

■　自然な共役分布

尤度を掛けることにより事前分布が同じ種類の事後分布に変換されるとき，その「事前分布と尤度を自然な共役分布」という．また，「自然共役な事前分布」と呼ぶ．

一般に，二項分布に従うデータから得られる尤度に対して，ベータ分布は相性が良い．
事前分布（ベータ分布）→尤度（二項分布）→事後分布（ベータ分布）

θがベータ分布$Be(\alpha, \beta)$に従う

確率分布：$f(\theta) = $ 定数 $\times \theta^{\alpha-1}(1-\theta)^{\beta-1}$

事前分布$\pi(\theta) = 1$はベータ分布の$\alpha = \beta = 1$の場合

ベイズ統計において，モデルから得られる尤度に対して自然な共役分布を用いることがポイントになる．

最終試験
予想問題と解説

- ➢試験範囲：全範囲
- ➢持ち込み不可

- ➢予想問題
- 授業中に行った演習問題
- 小テスト（２回分）及びその予想問題
- 今回の予想問題（１月２３日）

★試験問題は上記の予想問題から出題されます．
数値は若干変更される可能性あり．

問題1＜漸化式＞

次の漸化式の一般解を求めよ.

$$a_n = 4a_{n-1} - 3a_{n-2}, n \geq 2$$
$$a_0 = 0, a_1 = 1$$

＜解答例＞

次のように変形する（a_n～a_{n-2}を左辺に集める）と右辺＝0であり，同次解が一般解となる.

$$a_n - 4a_{n-1} + 3a_{n-2} = 0$$

特性方程式

$a_n = K\alpha^n$と仮定して，漸化式に代入し特性方程式を得る.

$$K\alpha^n = 4K\alpha^{n-1} - 3K\alpha^{n-2}$$
$$\alpha^2 - 4\alpha + 3 = 0 \quad \cdots \text{特性方程式}$$
$$\alpha = 3, 1$$

一般解（未定係数）

$$a_n = K_1 3^n + K_2 1^n$$

境界条件

$$a_0 = K_1 3^0 + K_2 1^0 = K_1 + K_2 = 0$$
$$a_1 = K_1 3^1 + K_2 1^1 = 3K_1 + K_2 = 1$$

より，

$$K_1 =, \frac{1}{2}\ K_2 = -\frac{1}{2}$$

一般解（最終）

$$a_n = \frac{1}{2}3^n - \frac{1}{2}1^n = \frac{1}{2}3^n - \frac{1}{2}$$

問題２＜漸化式＞
次の漸化式の一般解を求めよ.

$$a_n - a_{n-1} - 6a_{n-2} = 5$$
$$a_0 = 1, a_1 = 2$$

＜解答例＞
この漸化式（a_n~a_{n-2}を左辺に集める）では右辺$\neq 0$であるので，一般解＝同次解＋特解となる.

先ず，同次解を求める．漸化式の右辺＝0とする.

$$a_n - a_{n-1} - 6a_{n-2} = 0$$

特性方程式
$a_n = K\alpha^n$とおき，両辺を$K\alpha^{n-2}$で割る.

$$\alpha^2 - \alpha - 6 = 0$$
$$\alpha = 3, -2$$

同次解

$$a_n = A_1 \cdot 3^n + A_2 \cdot (-2)^n$$

次に，特解を求める．

$$a_n - a_{n-1} - 6a_{n-2} = 5$$

を満たす一つの解を求める．右辺が定数であるから

$a_n = Bn + C$と仮定して，漸化式に代入する．

$$a_n - a_{n-1} - 6a_{n-2} = (Bn + C) - \{B(n-1) + C\}$$
$$-6\{B(n-2) + C\} = B(-6n + 13) - 6C = 5$$

これより，

$$B = 0, \qquad C = -5/6$$

以上より，特解は次のようになる．

$$a_n = -5/6$$

一般解(未定係数)＝同次解＋特解

$$\alpha_n = A_1 \cdot 3^n + A_2 \cdot (-2)^n - 5/6$$

境界条件よりA_1, A_2を決める.

$$\alpha_0 = A_1 \cdot 3^0 + A_2 \cdot (-2)^0 - 5/6 = 1$$
$$\alpha_1 = A_1 \cdot 3^1 + A_2 \cdot (-2)^1 - 5/6 = 2$$

式を整理する.

$$\alpha_0 = A_1 + A_2 = 11/6$$
$$\alpha_1 = 3A_1 - 2A_2 = 17/6$$

より, $A_1 = 13/10, \quad A_2 = 8/15$

一般解(最終)

$$\alpha_n = \frac{13}{10} 3^n + \frac{8}{15} (-2)^n - 5/6$$

問題3＜漸化式＞

次の漸化式の一般解を求めよ.

$$a_n - 5a_{n-1} + 4a_{n-2} = 2^n, \qquad n \geq 2$$
$$a_0 = 1, \qquad a_1 = 4$$

＜解答例＞

まず，同次解を求める．漸化式の右辺＝0とする.

$$a_n - 5a_{n-1} + 4a_{n-2} = 0$$

特性方程式

$a_n = A\alpha^n$として，両辺を$A\alpha^{n-2}$で割る.

$$\alpha^2 - 5\alpha + 4 = 0$$
$$\alpha = 4, 1$$

同次解

$$a_n = A_1 \cdot 4^n + A_2 \cdot 1^n = A_1 \cdot 4^n + A_2$$

次に，特解を求める．

漸化式の右辺が2^nであることを考慮して，特解を$a_n = B \cdot 2^n$と仮定し，これが漸化式を満たすようにBを決める．

$$\begin{gathered} a_n - 5a_{n-1} + 4a_{n-2} \\ = B \cdot 2^n - 5B \cdot 2^{n-1} + 4B \cdot 2^{n-2} \\ = B \cdot 2^{n-2}(4 - 10 + 4) = -B \cdot 2^{n-1} = 2^n \end{gathered}$$

これより，

$$B = -2$$

特解

$$a_n = -2 \cdot 2^n = -2^{n+1}$$

一般解（未定係数）＝同次解＋特解

$$a_n = A_1 \cdot 4^n + A_2 - 2^{n+1}$$

境界条件より

$$a_0 = A_1 \cdot 4^0 + A_2 - 2^{0+1} = 1$$
$$a_1 = A_1 \cdot 4^1 + A_2 - 2^{1+1} = 4$$

式を整理する.

$$a_0 = A_1 + A_2 = 3$$
$$a_1 = 4A_1 + A_2 = 8$$

より

$$A_1 = \frac{5}{3}, A_2 = \frac{4}{3}$$

一般解（最終）

$$a_n = \frac{5}{3} 4^n + \frac{4}{3} - 2^{n+1}$$

問題4＜漸化式＞

次の漸化式の一般解を求めよ.

$$f_n - 2f_{n-1} = n - 3, \qquad n \geq 4$$

$$f_3 = 1$$

＜解答例＞

まず，同次解を求める. 漸化式の右辺＝0とする.

$$f_n - 2f_{n-1} = 0,$$

特性方程式

$f_n = K\alpha^n$を代入して，両辺をKf^{n-1}で割る.

$$K\alpha^n - 2K\alpha^{n-1} = 0$$

$$\alpha - 2 = 0 \rightarrow \alpha = 2$$

同次解

$$f_n = K \cdot 2^n$$

次に，特解を求める.

漸化式の右辺＝（nの１次式）であることを考慮して，

$f_n = An^2 + Bn + C$と仮定し，これが漸化式を満たすようにA, B, Cを決める.

$$
\begin{aligned}
&f_n - 2f_{n-1} \\
&= (An^2 + Bn + C) - 2\{A(n-1)^2 + B(n-1) + C\} \\
&= A(-n^2 + 4n - 2) + B(-n + 2) - C = n - 3
\end{aligned}
$$

上式において，

n^2の係数比較：$-An^2 = 0$より$A = 0$

nの係数比較：$(4A - B)n = n$より$B = -1$

定数項の比較：$2B - C = -3$より$C = 1$

特解

$$f_n = -n + 1$$

一般解(未定係数)＝同次解＋特解

$$f_n = K \cdot 2^n - n + 1$$

境界条件

$$f_3 = K \cdot 2^3 - 3 + 1 = 1$$

より

$$K = \frac{3}{8}$$

一般解(最終)

$$f_n = \frac{3}{8} 2^n - n + 1$$

漸化式の問題に関しては[例4.3], [例4.9], [例4.10], [例4.11]も予想問題になります.

問題5＜ベイズの定理＞

４人の死刑囚A，B，C，Dの中で1人だけ無作為に選ばれ，恩赦を受ける．誰が恩赦になるか看守は知っているが，囚人は知らない．

囚人Aが看守に「B，C，Dのうち誰かは必ず処刑されるから，その人の名前を教えてほしい」と頼んだ．看守はもっともだと思って「囚人Bが処刑される」と教えた．

囚人Aは「はじめは助かる確率は1/4であった」が，情報を得てから「助かるのは自分かC，Dなので，確率は1/3に上がった」と喜んだ．囚人Aの計算は正しいか？

（ベイズの定理により囚人Aが助かる確率を計算）

教科書（p.57）の同じ問題も予想問題になります．

事象を決める

事象A　Aが助かる

事象B　Bが助かる

事象C　Cが助かる

事象D　Dが助かる

事象SB　看守が「Bが処刑される」とAに教える

求める確率　$P(A|S_B)$

$$P(A|S_B) = \frac{P(S_B|A)P(A)}{P(S_B|A)P(A) + P(S_B|B)P(B) + P(S_B|C)P(C) + P(S_B|D)P(D)}$$

- 恩赦を受ける囚人は無作為に選ばれる（事前情報無し）

$$P(A)=P(B)=P(C)=P(D)=1/4\ (*)$$

- Bが助かる場合：「Bが処刑される」とは言わない．

$$P(S_B|B)=0$$

- Aが助かる場合：B，C，Dのうち誰かが処刑されるので，

$$P(S_B|A)=1/3$$

- CあるいはDが助かる場合：B，C，Dのなかで処刑されるのは「BまたはD」あるいは「BまたはC」であるから

$$P(S_B|C)=P(S_B|D)=1/2$$

以上より，

$$P(A|S_B)=\frac{\frac{1}{3}\times\frac{1}{4}}{\frac{1}{3}\times\frac{1}{4}+0\times\frac{1}{4}+\frac{1}{2}\times\frac{1}{4}+\frac{1}{2}\times\frac{1}{4}}=\frac{1}{4}\ (*)$$

となり，囚人Aの計算は間違っている．

問題6＜ベイズの定理＞

四つのドア（A，B，C，D）があり，どれかに賞金が隠されている．回答者が一つのドア（A）を選んだ．出題者が残りのドアから，はずれのドア（C）を開けた．

回答者は①ドアAのままにする，②ドアBかDを選び直すという2通りを選択できる．①，②のどちらが賞金を獲得する確率が高いか？
各々の確率を計算して比較すること．
（ベイズの定理により解答すること）

教科書にある同じ問題（p.61の例）も予想問題になります．

＜解答例＞

事象W　ドアAが当たり
事象X　ドアBが当たり
事象Y　ドアDが当たり
事象Z　ドアCを開く

求める確率

$P(W|Z), P(X|Z), P(Y|Z)$　いずれが高くなるかを調べる

ベイズの定理より

$$P(W|Z) = \frac{P(Z|W)P(W)}{P(Z|W)P(W) + P(Z|X)P(X) + P(Z|Y)P(Y)}$$

$$P(X|Z) = \frac{P(Z|X)P(X)}{P(Z|W)P(W) + P(Z|X)P(X) + P(Z|Y)P(Y)}$$

$$P(Y|Z) = \frac{P(Z|Y)P(Y)}{P(Z|W)P(W) + P(Z|X)P(X) + P(Z|Y)P(Y)}$$

$P(W)=P(X)=P(Y)=1/4$・・・はじめは情報がない

$P(Z|W)=1/3$・・・Aが当たりであれば，はずれはB，C，Dであるから Cを開く確率は1/3

$P(Z|X)=1/2$・・・ Bが当たりであれば，はずれはCかDであるから，Cを開く確率は1/2

$P(Z|Y)=1/2$・・・ Dが当たりであれば，はずれはBかCであるから，Cを開く確率は1/2

$$P(W|Z)=\frac{\frac{1}{3}\times\frac{1}{4}}{\frac{1}{3}\times\frac{1}{4}+\frac{1}{2}\times\frac{1}{4}+\frac{1}{2}\times\frac{1}{4}}=\frac{2}{8}\ (*)$$

$$P(X|Z)=P(Y|Z)=\frac{\frac{1}{2}\times\frac{1}{4}}{\frac{1}{3}\times\frac{1}{4}+\frac{1}{2}\times\frac{1}{4}+\frac{1}{2}\times\frac{1}{4}}=\frac{3}{8}\ (*)$$

$P(W|Z)<P(X|Z)=P(Y|Z)$より，ドアBまたはDを選び直すほうが当たる確率が高い．

問題7＜ベイズの定理＞

１個の壺がある．壺の中には白と赤の5個の玉が入っている．そこから玉１個を取りだしたとき，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．但し，壺の中にある赤玉の個数は奇数であることが分かっている．

教科書にある同じ問題（p.79の例）も予想問題になります．

＜解答例＞

仮定(原因)：壺の中の赤玉の個数＝1，3，5個

壺1[○○○○●]，壺2[○○●●●]，

壺3[●●●●●]

H_i：壺iから玉1個取り出す．$i = 1, 2, 3$

結果：取りだした玉が赤玉である．

D：壺から玉1個を取りだしたとき，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．

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

$P(D|H_i)$:壺iから赤玉を1個取り出す確率.

$$P(D|H_1) = 1/5, P(D|H_2) = 3/5, P(D|H_3) = 5/5$$

$P(H_i)$: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.

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

上記の確率を用いて目的の確率分布が求まる.

$$P(H_1|D) = \frac{\frac{1}{5}\times\frac{1}{3}}{\frac{1}{5}\times\frac{1}{3}+\frac{3}{5}\times\frac{1}{3}+\frac{5}{5}\times\frac{1}{3}} = \frac{1}{9}$$

赤玉の個数	1個	3個	5個
確率	$P(H_1|D) = 1/9$	$P(H_2|D) = 3/9$	$P(H_3|D) = 5/9$

確率の合計: $1/9 + 3/9 + 5/9 = 1$

問題8＜事前分布，尤度，事後分布＞

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2回投げたとき，１回目に表，２回目に裏が出た．このとき，表の出る確率θの事後分布に関して以下の問に答えよ．

１．事後分布の式を求めよ．
２．事後分布の概略図を描け．
３．$0.5 \leq \theta \leq 1$に対する確率を求めよ．

＜解答例＞

■対象となる母数：表の出る確率$= \theta, 0 \leq \theta \leq 1$

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

$$f(表|\theta) = \theta$$
$$f(裏|\theta) = 1 - \theta$$

■事前分布：$\pi(\theta)$ $\rightarrow \pi_0(\theta)$･･コインを投げる前の事前分布
「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情報はない．
「理由不十分の原則」に基づいて「一様分布」する．

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

■「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,

$$\pi_1(\theta) = \pi(\theta|D_1) = 2\theta$$

２回目のコイン投げに対する事前分布となる.

■「２回目に裏が出た」というデータを取り込む

D_2：2回目に裏が出る.

コインを2回投げた後のθの事後分布

$$\pi(\theta|D_2) \propto f(D_2|\theta) \times \pi_1(\theta) = (1-\theta) \times 2\theta = 2\theta(1-\theta)$$

規格化条件（面積＝1）より

$$\pi_2(\theta) = \pi(\theta|D_2) = 6\theta(1-\theta)$$

（答え）

１．事後分布の式

$$\pi_2(\theta) = \pi(\theta|D_2) = 6\theta(1-\theta)$$

２．事後分布の概略図　　板書する

３． $0.5 \le \theta \le 1$に対する確率

$$P(0.5 \le \theta \le 1) = \int_{0.5}^{1} 6\theta(1-\theta)d\theta = \frac{1}{2}$$

問題9＜事前分布，尤度，事後分布＞

３人の治験者を抽出し，新薬の効用を調べたところ，２人には効き，１人には効かないことが分かった．新薬の効き具合の分布を調べ，以下の問に答えよ.

１．新薬の効き具合の分布を式で表せ.
２．分布の概略図を示せ.
３．$0.5 \leq \theta \leq 1$に対する確率を求めよ.
４．平均値を求めよ.

＜解説＞

■ベイズ統計の母数

抽出した1人の治験者に新薬が効く確率θ

■尤度：$f(D|\theta)$

「効く確率」θのもとで，データD（3人のうち2人に効き，1人に効かない）の起こる確率（二項分布より求まる）

$$f(D|\theta) = {}_3C_2\theta^2(1-\theta)^1, 0 \le \theta \le 1$$

■事前分布：$\pi(\theta)$

治験するまでは効く確率θに関する情報はないので，理由不十分の原則より全ての可能性は均等であるとする．

$$\pi(\theta) = k$$

$0 \le \theta \le 1$であり，確率の総和（面積）＝1より，

$$\text{事前分布}\quad \pi(\theta) = 1, 0 \le \theta \le 1$$

■事後分布:$\pi(\theta|D)$

ベイズ統計の基本公式より

事後分布∝尤度×事前分布

$$\pi(\theta|D) \propto {}_3C_2\theta^2(1-\theta)^1 \times 1 \propto \theta^2(1-\theta)$$

$\pi(\theta|D)$の積分＝１より,

$$\int_0^1 k\theta^2(1-\theta)d\theta = \frac{k}{12} = 1 \rightarrow k = 12$$

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

（答え）

１．分布の式　$\pi(\theta|D) = 12\theta^2(1-\theta)$

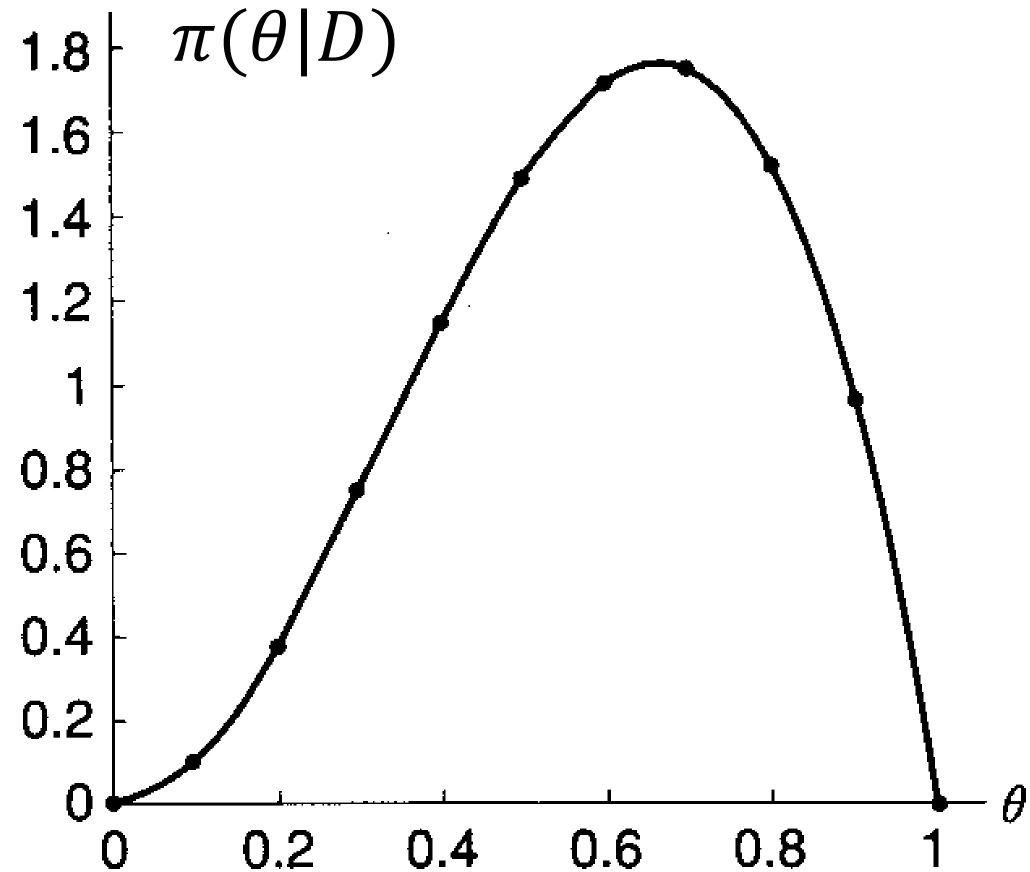
２．分布の概略図　次頁

３．$0.5 \leq \theta \leq 1$に対する確率

$$P(0.5 \leq \theta \leq 1) = \int_{0.5}^{1} 12\theta^2(1-\theta)d\theta = \frac{11}{16}$$

４．平均値

$$\mu = \int_0^1 \theta \times 12\theta^2(1-\theta)d\theta = \frac{3}{5}$$


$\pi(\theta|D)$
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
0.2
0.4
0.6
0.8
1
θ