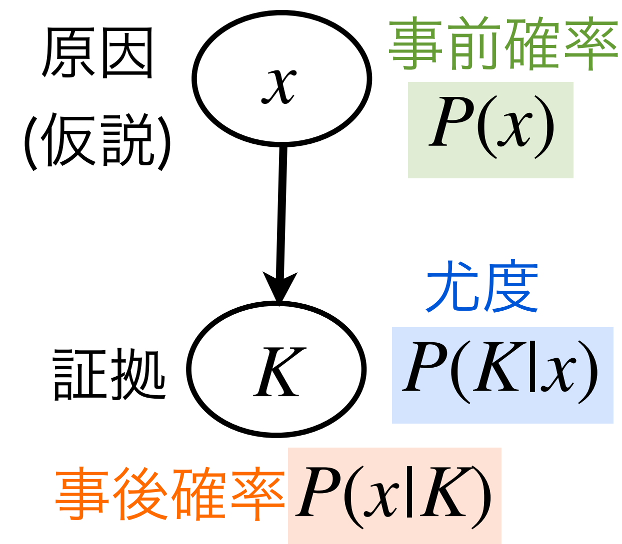
最尤推定とベイズ推定

血液型	各血液型の割合	几帳面(K)
O	$P(O)=0.4$	$P(K \mid O)=0.4$
A	$P(A)=0.3$	$P(K \mid A)=0.5$
B	$P(B)=0.2$	$P(K \mid B)=0.6$
AB	$P(AB)=0.1$	$P(K \mid AB)=0.5$



方法1) **ベイズ推定**:

$\boxed{}$ を使う!

$$血液型 = \arg\max_x P(x|K)$$

ベイズ推定の解: O型!

血液型の分布
(事前確率)を利用

?

方法2) **最尤推定**:

$\boxed{}$ を使う!

$$血液型 = \arg\max_x P(K|x)$$

最尤値: B型!

血液型の分布
(事前確率)は無視

$$P(x|K) = \frac{P(x)}{P(K)} P(K|x)$$

事前確率が不明なら両者の値は一致

ベイズ推定と最尤推定の例題

準備

確率と確率密度

1)さいころを振った時に3が出る確率は?

2) 0以上1以下の実数をランダムに発生する装置がある。

a) 0.5が出る確率は?

b) 0から0.5の間の数が出る確率は?

c) 0から1の間の数が出る確率は?

表の出る確率が全く未知の、いびつな形のコインがある。このコインを何度か振って，表が出る確率の真の値 x を推定したい。

事前知識が全くない時

$\rightarrow$ 確率 x は0から1のあらゆる値を取り得る

(全ての可能性は均等:　　　　　　　　の原則)

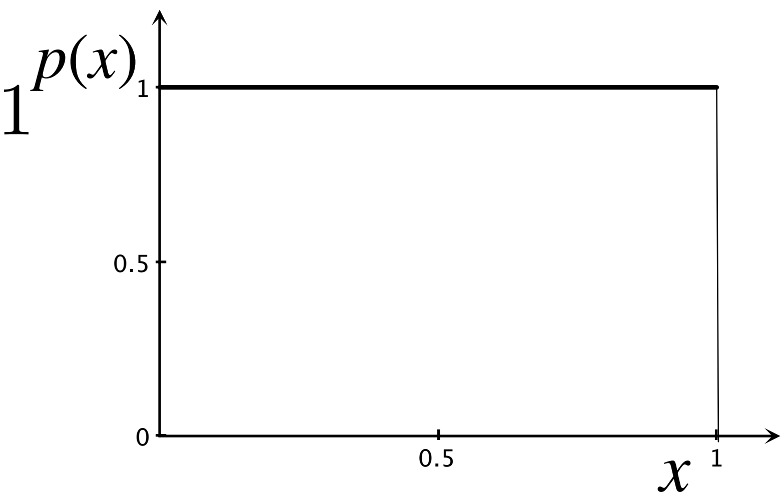
$\rightarrow$ 確率 x となる可能性(事前分布)は $p(x)=C$　(Cは定数)！

ただし，

$$P(0 \leq x \leq 1) = \int_0^1 p(x)dx = 1$$

より　　　$p(x) = 1$

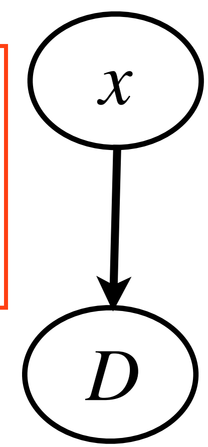
　　　　関数



最尤推定: コイン投げ

1回めは表が出た。表が出る確率の真の値xは?

(1) 確率xに対してD={表}が出る確率 (尤度) $L(x)$は?



事前分布:
$p(x)=1$

尤度:
$L(x)=P(D|x)$

(2) xの最尤値は？

最尤推定: コイン投げ

2回めは裏!

(1) 確率xに対してD={表, 裏}が出る確率(尤度)$L(x)$は?

(2) pの最尤値は？

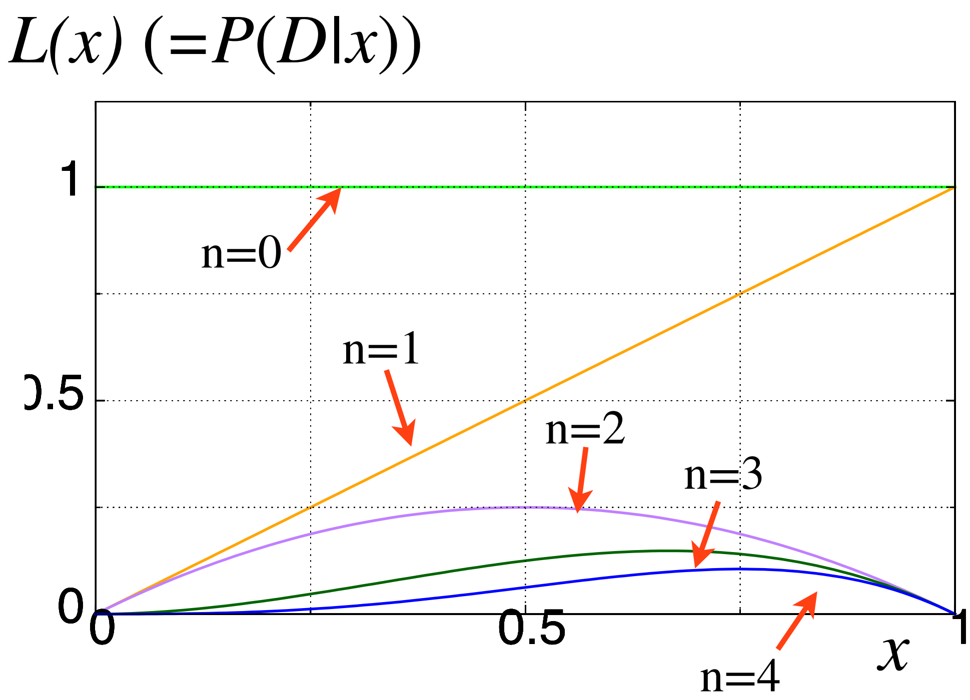
最尤推定: コイン投げ

4回振ったら {表, 裏, 表, 表}

(1) 確率xに対してD={表, 裏, 表, 表}が出る尤度$L(x)$は?

(2) xの最尤値は？

最尤推定のまとめ



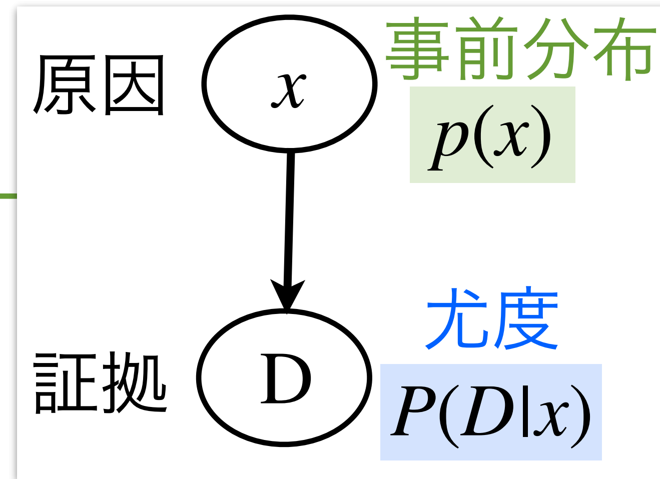
- データ数が増えていくと，尤度は(多くの場合)先鋭化
- 最頻値を最尤値と呼ぶ

ベイズ推定

ベイズ推定: コイン投げ

(1) データD={表}に基づいて、確率xのとりうる各値に対する自信を定量的に表しなさい。

事後分布 $p(x|D)$を求める!



ベイズ統計の基本式

$$p(x|D) = \frac{P(D|x)p(x)}{P(D)}$$

$$= kP(D|x)p(x)$$

ただし $\int_0^1 p(x|D)dx = 1$

ベイズ推定: コイン投げ

(1) データD={表}に基づいて、確率xのとりうる各値に対する自信を定量的に表しなさい。(事後分布 $p(x|D)$ を求める!)

(2) $0<x<1/2$ である確率は?

(3) 最も自信を持てるxの値は?

コインを4回投げたら，表，裏，表，表の順で出た。コインを投げる回数とともに，確率xのとりうる各値に対する自信(事後分布)はどう変化したか?

ベイズ推定: コイン投げ

2回投げた後 (D={表,裏}) の事後分布は？

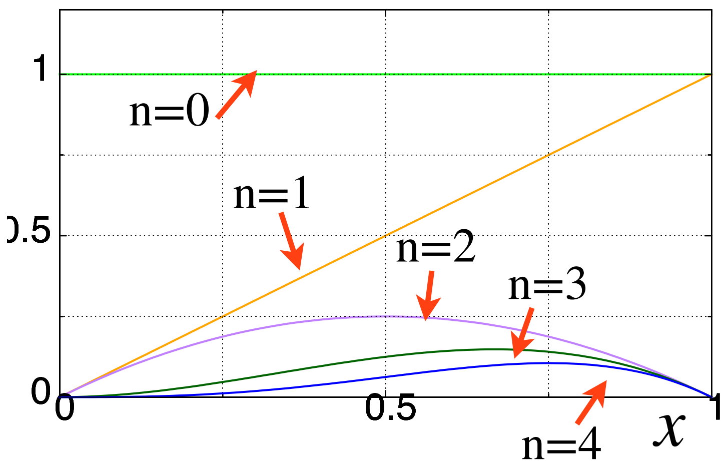
3回投げた後 (D={表,裏,表}) の事後分布は？

4回投げた後 (D={表,裏,表,表}) の事後分布は？

最尤推定とベイズ推定

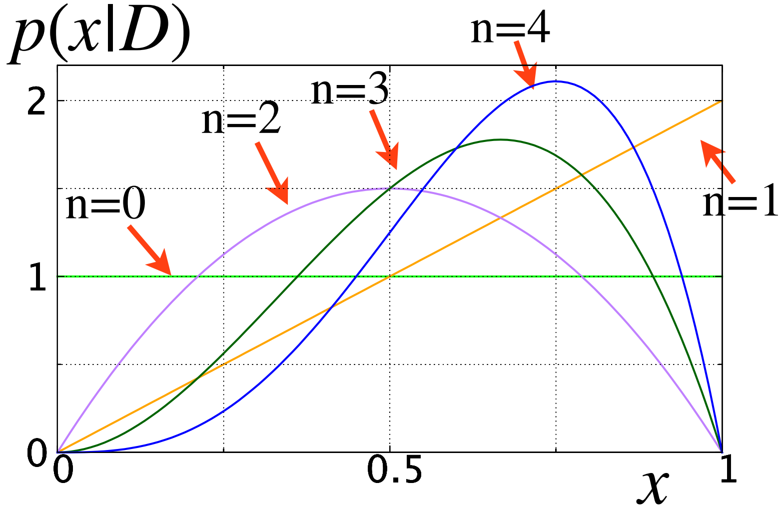
尤度

$P(D|x)$



事後分布

$p(x|D)$



最頻値: 最尤値

最頻値: ベイズ推定の解

- データ数とともに，尤度も事後分布も先鋭化
- 事前分布が不明なら最尤値とベイズ推定の解は一致
- (ヒトは最尤推定をしているらしい...)

ベイズ統計学入門

ある病気の新薬を開発し，その効果を試した。

(1) 3人の患者に試したところ，2人には効果が認められた。新規の患者に対して新薬の効果が認められる確率の分布を調べなさい

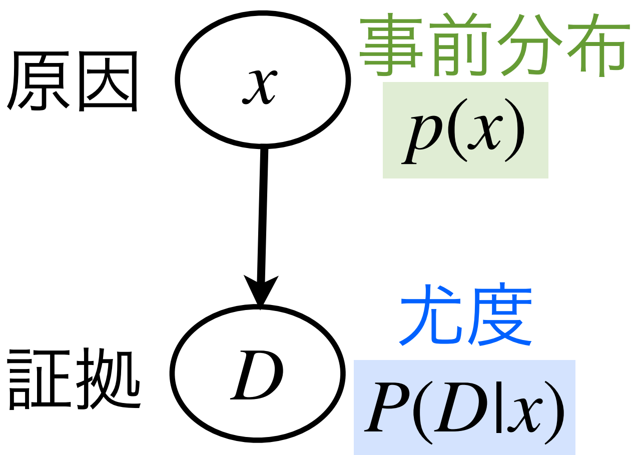
(2) 30人の患者に試したところ，20人には効果が認められた場合も同様に効果の分布を調べなさい。

(3)この新薬は少なくとも二人に一人に効果がある，と言える確率を求めなさい。

(1) 3人の患者に試したところ，2人には効果が認められた。新規の患者に対する，この新薬の効果の分布を調べなさい

データDに基づいて，新薬の効果 x

(何割のヒトに効果があるか)の分布$p(x|D)$を推定する



事前分布: $p(x) = 1$ (理由不十分の原理)

尤度: $P(D|x) = {}_3C_2 x^2(1-x)$

事後確率: $p(x|D) = \alpha p(x) P(D|x)$

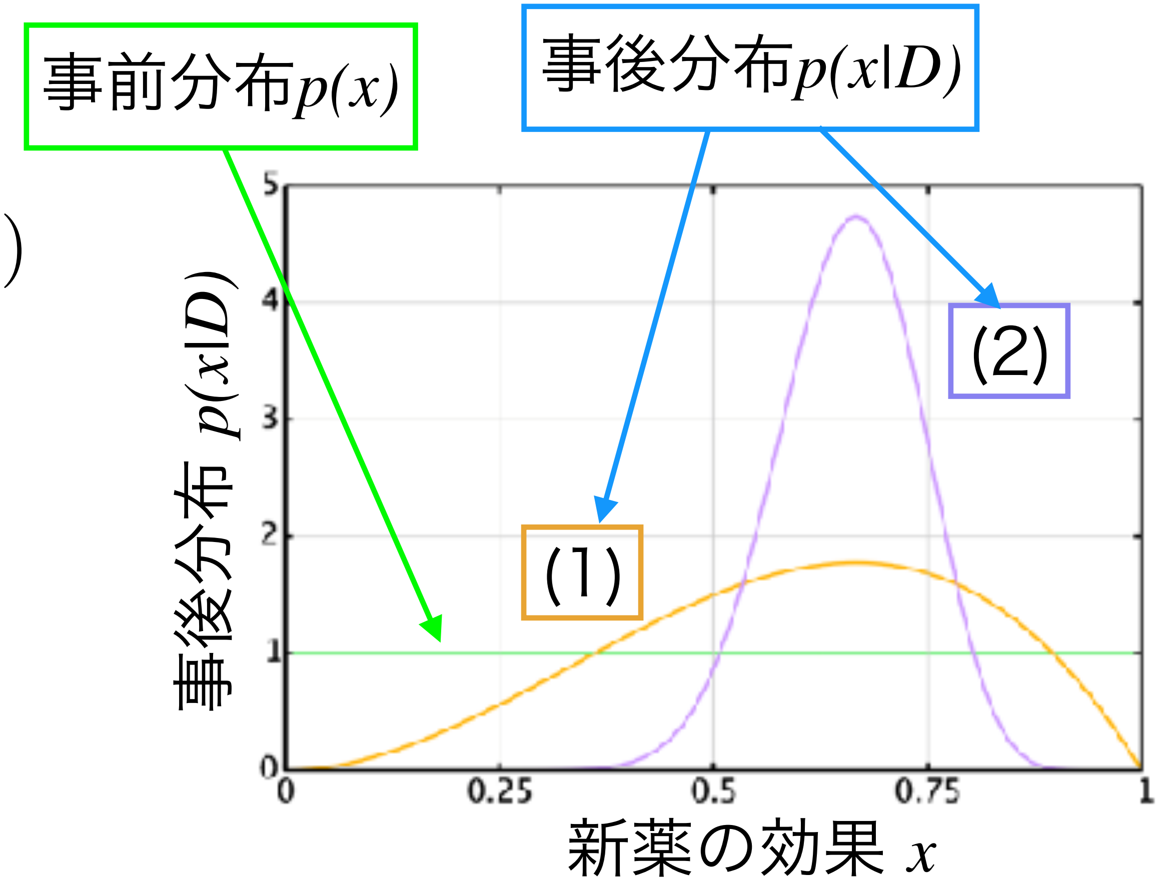
$$= \alpha \, {}_3C_2 x^2 (1-x)$$

x: 薬の効果(効く確率)

D: データ(何人中何人に効いた)

(1) 3人中2人に効果:

$$p(x|D) = 12x^2(1-x)$$



(2) 30人中20人に効果:

$$p(x|D) = 931395465x^{20}(1-x)^{10}$$