第2章　ベイズ決定の基礎

§2.1　ベイズ決定
§2.2　多次元ベイズの決定

05T4074A
久保田　敦

§2.1　ベイズ決定

➢ 損失関数による統計的決定

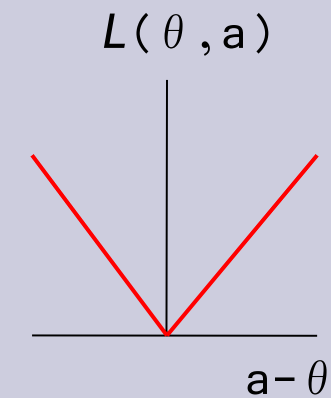
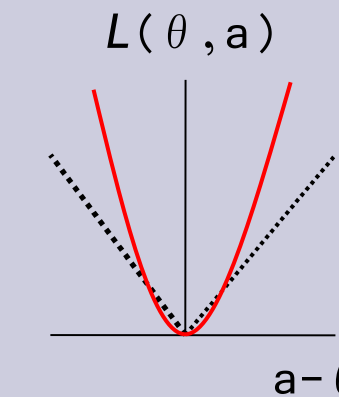
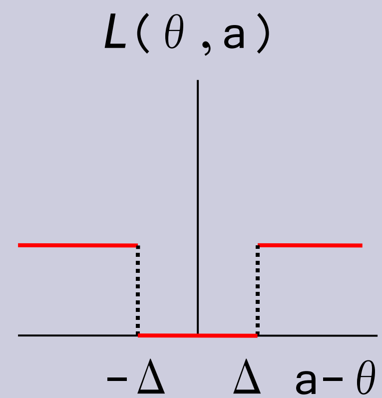
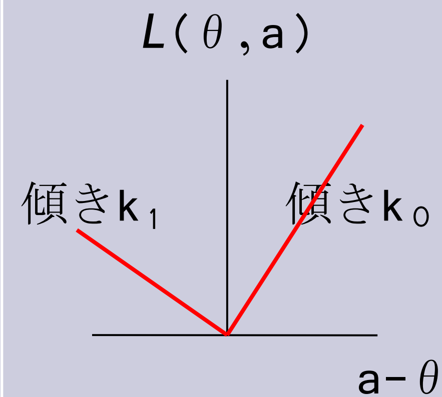
事後確率分布は θ の分布を示し、θ のありかを予想する。
推定問題とは1つの値を決定すること。

事後確率分布から1つの値を推定値として取り出す方法を考える

損失：推定の正確さを表す。推定値aが θ と離れるほど大きい

損失関数$L(a, \theta)$を定めてより損失がより小さいaを求める。
これを経済的決定理論という。

損失

損失	絶対損失	平方損失	0-1型単純損失	非対称絶対損失
式	$L(\theta,a)=\lvert a-\theta\rvert$	$L(\theta,a)=(a-\theta)^2$	$L(\theta,a)=\begin{cases}0, & \lvert a-\theta\rvert\leqq\Delta\\ 0, & \lvert a-\theta\rvert>\Delta\end{cases}$	$L(\theta,a)=\begin{cases}k_0\cdot\lvert a-\theta\rvert, & a\geqq\theta\\ k_1\cdot\lvert a-\theta\rvert, & a<\theta\end{cases}$
説明	誤差に正しく比例した損失	大きい誤差に対して比例以上に制裁	誤差Δ内の「あたり」か、「はずれ」のどちらか	過大推定と過小推定でペナルティーが異なる
グラフ				

➢リスク

各データ z をそれに最適な行動aに対応させる関数

$$R(\theta, d) \equiv \sum_{z} p(z \mid \theta) \cdot L(\theta, d(z))$$

θ についてさらに平均をとる

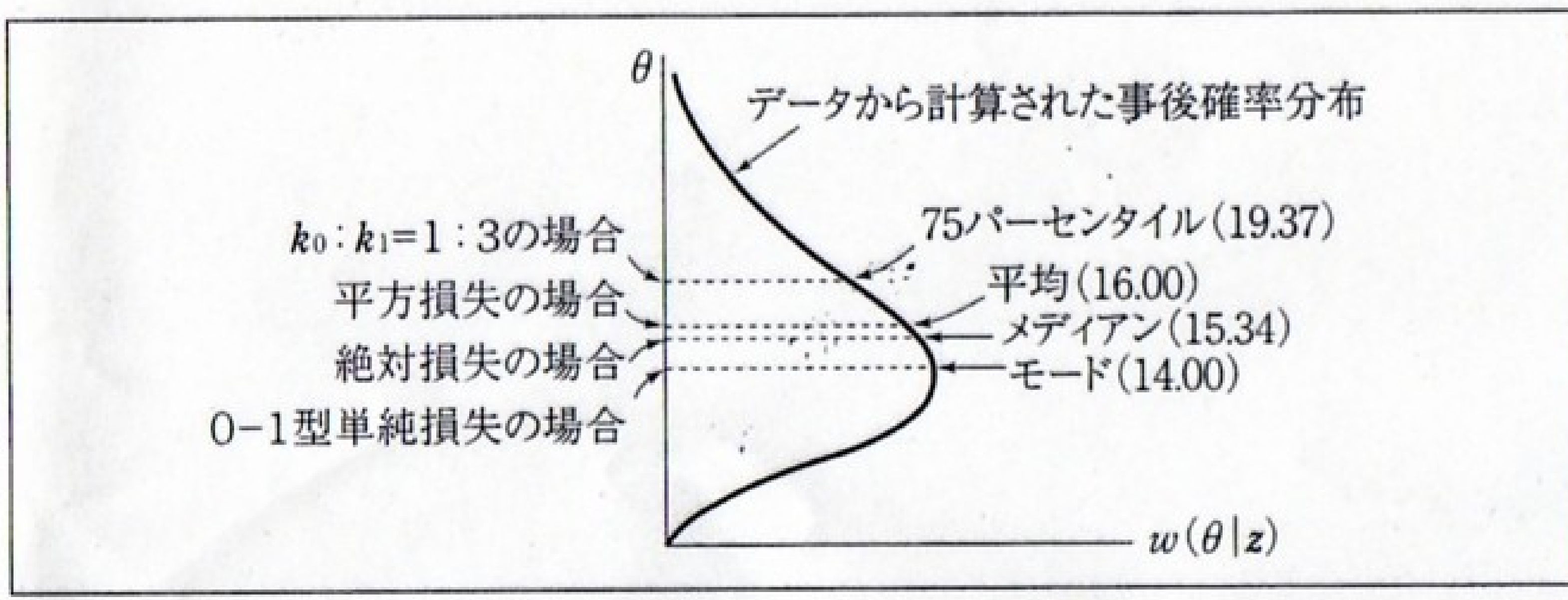
$$\begin{aligned} r(d) &\equiv \sum_{z} w(\theta) R(\theta, d) \\ &= \sum_{\theta} \sum_{z} w(\theta) \cdot p(z \mid \theta) \cdot L(\theta, d(z)) \\ &= \sum_{z} p(z) \left\{ \sum_{\theta} w'(\theta) \cdot L(\theta, d(z)) \right\} \end{aligned}$$

- r(d)を最小にするdを求めることをベイズ決定
- 事後期待損失{　}を最小化することで求められる。

ベイズ決定

損失	問題	解
絶対損失	$\underset{a}{Min}\sum_{\theta} w'(\theta \mid z)\cdot\lvert\theta - a\rvert$	中央値
平方損失	$\underset{a}{Min}\sum_{\theta} w'(\theta \mid z)\cdot(\theta - a)^2$	平均値
0-1型単純損失	$\underset{a}{Min}\sum_{\theta} w'(\theta \mid z)\cdot(1 - \overline{L}(\theta, a))$	モード（最頻値）
非対称絶対損失	$100\cdot\dfrac{k_1}{k_0 + k_1} = \alpha$	100αパーセンタイル

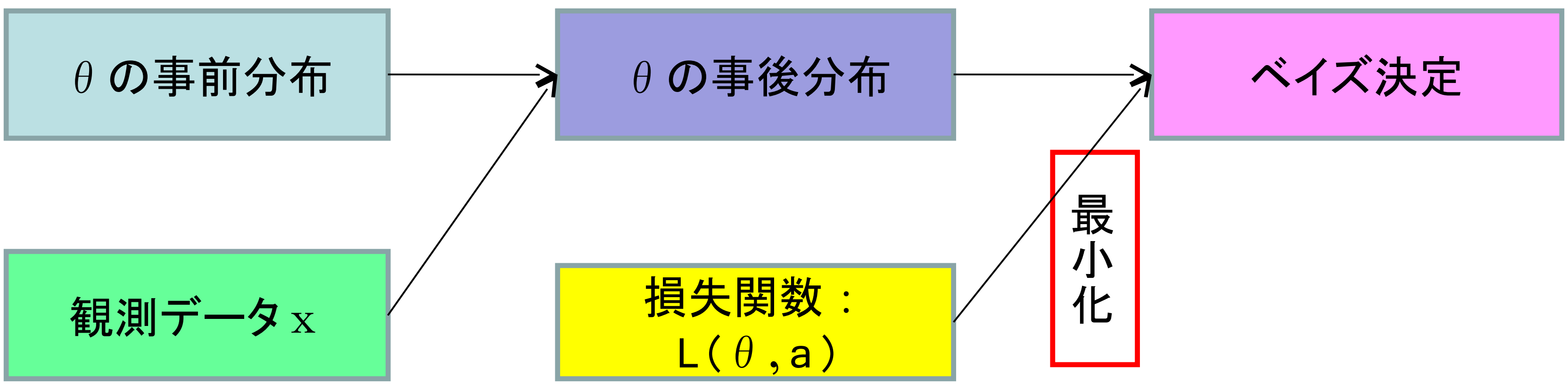
ベイズ決定の例


θ
データから計算された事後確率分布
75パーセンタイル(19.37)
平均(16.00)
メディアン(15.34)
モード(14.00)
k0：k1=1：3の場合
平方損失の場合
絶対損失の場合
0−1型単純損失の場合
w(θ|z)

§2.2　多次元のベイズ決定

ベイズ決定を様々な分野のシステムとして用いたい場合には多次元の要素を扱うことも必要である。

➢<u>統計的決定の基本原理</u>



ベイズ決定は全てこの流れの応用である
。

➢ 多次元正規分布の平均推定

確立変数 $X^{(1)},\cdots,X^{(m)}$ はp次元正規分布：$N(\mu\ ,\Sigma\ )$に従う
（**μ**　：平均ベクトル(未知)、**Σ**　：分散・共分散行列(既知)）

決定者の「行動」**a**は μ を推定するp次元ベクトル

損失関数：L(θ , a) $= (\boldsymbol{\mu} - \mathbf{a})'\mathbf{M}(\boldsymbol{\mu} - \mathbf{a})$
（Mは(P× P)の正値定符号、対称行列）

μ (θ)の事前分布 $= N(\lambda\ ,T)$ (p次元正規分布)

標本平均ベクトルm $= \dfrac{x^{(1)} + \cdots + x^{(n)}}{n}$

μ の事前分布 $N(\lambda\ ,T)$ $\xrightarrow{\ m\ }$ μ の事後分布 $N(\tilde{\lambda}\ ,\tilde{T})$

➢ 多次元正規分布の平均推定2

$$\tilde{\lambda} = \tilde{T}(T^{-1}\lambda + n\sum{}^{-1} m), \tilde{T} = (T^{-1} + n\sum{}^{-1})^{-1}$$

これは、正規分布のベイズ更新の一般化。

よって $(\mu - a)' M(\mu - a)$ の $N(\tilde{\lambda}, \tilde{T})$

に関する期待値を最小化するaを求めればベイズ決定方式が求まる

$$E(\mu - a)' M(\mu - a)$$

$$= E(\mu - \tilde{\lambda})' M(\mu - \tilde{\lambda}) + (\tilde{\lambda} - a)' M(\tilde{\lambda} - a)' M(\tilde{\lambda} - a)$$

この左辺の最小値は $a = \tilde{\lambda}$ のとき

よってベイズ決定方式は

$$d_0(m) = \tilde{\lambda} = \tilde{T}(T^{-1}\lambda + nT^{-1}m)$$

➢ 変数の効き方

データが　m　から　$\Delta m = (\Delta m_0, \Delta m_1, \cdots, \Delta m_p)$

だけ変動したとする。それによるリスクr は

$$r = (\mathbf{\Delta}\tilde{\boldsymbol{\lambda}})' M \mathbf{\Delta}\boldsymbol{\lambda} = n^2 (\Delta m)' \Gamma (\Delta m)$$

$$\Gamma = (\tilde{\mathbf{T}}\Sigma^{-1})\mathbf{M}(\tilde{\mathbf{T}}\Sigma^{-1})$$

この Γ がr を各次元に分ける働きをする行列であり、
各次元のリスクへの効き方を示す。