No. 968

Bayes検定についての考察
-決定理論的アプローチ-

by

Yoshiko Nogami

January 2002

Bayes検定についての考察
ーー決定理論的アプローチーー

野　上　佳　子

概要：

決定理論を用いたBayes検定は、損失関数の設定が重要ある。この論文では、参考文献、鈴木[3]にもとずいて、これまでのBayes検定の方式 を復習し(§2)、片側検定の場合は、線型損失関数(例えば、 R. S. Singh[1]を、参照)を用いて、両側検定の場合は、積距離(product distance)損失関数(例えば、R. S. Singh & Wei Laisheng[2]を、参照)を用いてBayes解を求める。§5では、自然共役事前分布族を用いた例題を示す。

§1. 序章。

検定論には、R. A. Fisherの有意性検定と、それを発展させたJ. Neyman and E. S, Pearsonの仮説検定があげられる。この論文では、決定理論の立場から、Bayes検定という推論形式を論じる。

$X_1, \ldots, X_n$を、自然の状態（パラメータ）が、$\Theta=\theta$の時の確率（密度）関数$f(x|\theta)$からの大きさnのランダム標本とし、$\xi(\theta)$をΘの事前分布とする。§2 では、これまでのBayes検定の考え方について述べる。 §3 では、仮説 $H_0:\theta\leqq\theta_0$ 対 $H_1:\theta>\theta_0$ （θ_0:定数）とした時の決定問題を、線型損失関数(例えば、R. S. Singh[1]を参照）を用いてBayes解を考える。§4では、２つの仮説を$H_0:\theta_1\leqq\theta\leqq\theta_2$ 対 $H_1:\theta<\theta_1$ 或いは $\theta_2<\theta$ （θ_1 とθ_2は、$\theta_1\leqq\theta_2$なる実数）とした時の決定問題を、積距離(product distance)損失関数(例えば、R. S. Singh & Wei Laisheng[2]を参照）を用いて、Bayes解を考える。§5では、§3と§4の例を正規母集団と正規事前分布の例とベルヌイ母集団とベータ事前分布の例を用いて紹介する。

§2. Bayes 検定。

この節は、参考文献、鈴木[3]より、これまでのBayes検定の考え方について述べる。

決定問題を規定するものには、

(1) 決定の空間 D

(2) 自然の状態の空間（Θの空間）Ω

(3) 損失関数$l(d,\theta)$（或いは、効用関数$U(d,\theta)$）

(4) 確率変数Xの空間 $\mathfrak{X}$

(5) Xの確率（密度）関数 $f(x|\theta)$ 及び Θの事前分布$\xi(\theta)$

があげられる。ある決定問題に対して、上述の(1)~(5)が規定されたとして、この問題における最適決定を求める手続きを考える。

$X_1, \ldots, X_n$を、自然の状態$\Theta=\theta$の時の確率（密度）関数$f(x|\theta)$からの大きさnのランダム標本とする。$\underset{\sim}{x}=(x_1, \ldots, x_n)'$とする。実験結果が、$\underset{\sim}{X}=\underset{\sim}{x}$ の時、$\underset{\sim}{X}=\underset{\sim}{x}$が与えられた時の θ の事後確率（密度)関数は、

$$\xi(\theta|\underset{\sim}{X}=\underset{\sim}{x})=\xi(\theta|\underset{\sim}{x})= f(\underset{\sim}{x}|\theta)\xi(\theta)/ \int_{\Omega} f(\underset{\sim}{x}|\theta)\xi(\theta)\, d\theta$$

$$\propto f(\underset{\sim}{x}|\theta)\xi(\theta)$$

で与えられる。 充足統計量$T=t(X_1, \ldots, X_n)$が存在するならば、$g(t|\theta)$をその確率（密度）関数とする時、$T=t$ が与えられた時のθの事後確率（密度）関数は、

$$\xi(\theta|T=t)=\xi(\theta|t)=g(t|\theta)\xi(\theta)/\int_{\Omega} g(t|\theta)\xi(\theta)\, d\theta$$

$$\propto g(t|\theta)\xi(\theta)$$

で与えられる。 $t(\underset{\sim}{x})=t$ という関係が、$\underset{\sim}{x}$とtの間に存在すれば、$\xi(\theta|\underset{\sim}{x})=\xi(\theta|t)$が、成立する。

仮説 $H_0:\theta\in\Omega_0$ $(\Omega_0\subset\Omega)$を検定する問題を考える。実験結果が、$\underset{\sim}{X}=\underset{\sim}{x}$の時、この仮説が正しい確率は、$\xi(\theta|\underset{\sim}{x})$を事後確率（密度）関数とすると、

$$P(\theta\in\Omega_0|\underset{\sim}{X}=\underset{\sim}{x})=\int_{\Omega_0}\xi(\theta|\underset{\sim}{x})\,d\theta$$

で与えられる。 この時、

$$P(\theta\in\Omega_0|\underset{\sim}{X}=\underset{\sim}{x})=\begin{cases}1 & \text{ならば、}\quad H_0\text{は、正しい}\\ 0 & \text{ならば、}\quad H_0\text{は、正しくない}\end{cases}$$

と考える。

次の抜き取り検査の例を用いて、Bayes検定の考え方を説明する。

例） ある工程で作られるロットの不良率をθとすると、θがθ_0以下の時、そのロットを合格、θ_0より大であれば、不合格とする基準が定められているとする。同じ工程で作られた新しいロットを合格、不合格のいづれに判断するかを抜き取り検査の結果に基づいて行なう問題を考える。$\Omega_0=\{\theta:0\leqq\theta\leqq\theta_0\}$として、仮説 $H_0:\theta\in\Omega_0$ 対 対立仮説$H_1:\theta\in\Omega_0$ を考える。θの事前確率（密度）関数$\xi(\theta)$は、これまでの同工程による生産の過程で蓄積された知識と情報から想定されたものとする。抜き取り検査により、$\underset{\sim}{X}=\underset{\sim}{x}$が与えられた時、$H_0$と$H_1$が正しい確率は、それぞれ、

$$P(H_0|\underset{\sim}{x})=P(\theta\in\Omega_0|\underset{\sim}{X}=\underset{\sim}{x})$$

$$P(H_1|\underset{\sim}{x})=P(\theta\in\Omega_0|\underset{\sim}{X}=\underset{\sim}{x})=1-P(H_0|\underset{\sim}{x})$$

のように求まる。抜き取り検査を、行なう前は、それらは、$P(H_0)=P(\theta\in\Omega_0)$, $P(H_1)=P(\theta\notin\Omega_0)=1-P(H_0)$なので、$\underset{\sim}{X}=\underset{\sim}{x}$という標本情報により、$H_i\,(i=0,1)$が正しい確率は、$P(H_i)$から、$P(H_i|\underset{\sim}{x})$に変わる。

この$P(H_0|\underset{\sim}{x})$に基づいて、合否の判定を考える。$P(H_0|\underset{\sim}{x})\geqq P(H_1|\underset{\sim}{x})$、すなわち、$p(H_0|\underset{\sim}{x})\geqq 0.5$なら合格、そうでなけば、不合格という判断は、妥当と思われない。なぜなら、２種類の誤判別

(i) 本当は、合格であるべきロットを不合格とする

(ii)本当は、不合格であるべきロットを合格とする

が考えられ、両者の影響度、軽重の度合いは、異なるからである。 従って、損失の評価が、必要となる。H_iが正しいとする決定をd_iとし、決定の空間を、$D=\{d_0,d_1\}$とする。aとかbのように定数でθに依存しない損失関数$L(d_1,H_0)$、$L(d_0,H_1)$を考える。$\underset{\sim}{X}=\underset{\sim}{x}$を、得

た時、決定d_0をとる時の期待損失を

(2.1) $$R(d_0|\underset{\sim}{x})=L(d_0,H_1)P(H_1|\underset{\sim}{x})=L(d_0,H_1)(1-P(H_0|\underset{\sim}{x}))$$

$\underset{\sim}{X}=\underset{\sim}{x}$を得た時、決定$d_1$をとる時の期待損失を

$$R(d_1|\underset{\sim}{x})=L(d_1,H_0)P(H_0|\underset{\sim}{x})$$

とし、

$R(d_0|\underset{\sim}{x})\leqq R(d_1|\underset{\sim}{x})$ ならば、d_0を選択

$R(d_0|\underset{\sim}{x})>R(d_1|\underset{\sim}{x})$ ならば、d_1を選択

のように、とるべき決定を決めるのが合理的と考えられる。(2.1)式の最右辺より、上の決定方式は、

$L(d_0,H_1)/\{L(d_0,H_1)+L(d_1,H_0)\}\leqq p(H_0|\underset{\sim}{x})$ ならば、d_0を選択

$> p(H_0|\underset{\sim}{x})$ ならば、d_1を選択

とも書ける。

次に、仮説検定から離れて、この問題を純粋な決定問題と考える。決定d_0をロット合格、決定d_1をロット不合格とし、損失関数$l(d_i,\theta)(i=0,1)$が、Ω上で定義されているとする。$\underset{\sim}{X}=\underset{\sim}{x}$が与えられた時、条件付き期待 損失は、

$$R(d_i|\underset{\sim}{x})=E(l(d_i,\theta)|X=\underset{\sim}{x})=\int_\Omega l(d_i,\theta)\,\xi(\theta|\underset{\sim}{x})\,d\theta$$

$(i=0,1)$で与えられるので、

$R(d_0|\underset{\sim}{x})\leqq R(d_1|\underset{\sim}{x})$ ならば、d_0を選択

$R(d_0|\underset{\sim}{x})>R(d_1|\underset{\sim}{x})$ ならば、d_1を選択

によって、最適決定が決まる。

ここで特に、

$$l(d_0,\theta)=\begin{cases}L(d_0,H_1), & \theta\notin\Omega_0\\ 0, & \theta\in\Omega_0\end{cases}$$

$$l(d_1,\theta)=\begin{cases}0, & \theta\notin\Omega_0\\ L(d_1,H_0), & \theta\in\Omega_0\end{cases}$$

とすると、上述の例の仮説検定と一致する。

§3. 片側検定。

この節では、§2と同じく片側検定$H:\theta\leqq\theta_0$ 対 $H_1:\theta>\theta_0$を考える。$\Omega_0=(-\infty,\theta_0]$とする。ここでは、線型損失(R. S. Singh[1]を参照)を仮定する。すなわち、bを、b>0なる定数として、

$$l(d_0,\theta)=\begin{cases}b(\theta-\theta_0), & \theta\notin\Omega_0\\ 0, & \theta\in\Omega_0\end{cases}$$

$$l(d_1,\theta)=\begin{cases}0, & \theta\notin\Omega_0\\ b(\theta_0-\theta), & \theta\in\Omega_0\end{cases}$$

とする。事後確率（密度）関数 $\xi(\theta|\underset{\sim}{x})$を用いて、条件付期待損失は、それぞれ、

$$R(d_0|\underset{\sim}{x})=E(l(d_0,\theta)|\underset{\sim}{X}=\underset{\sim}{x})=b\int_{\theta_0}^{\infty}(\theta-\theta_0)\,\xi(\theta|\underset{\sim}{x})\,d\theta$$

$$R(d_1|\underset{\sim}{x})=E(l(d_1,\theta)|\underset{\sim}{X}=\underset{\sim}{x})=b\int_{-\infty}^{\theta_0}(\theta_0-\theta)\,\xi(\theta|\underset{\sim}{x})\,d\theta$$

と書けるので、これより

$$b^{-1}\{R(d_0|\underset{\sim}{x})-R(d_1|\underset{\sim}{x})\}=\int_{-\infty}^{\infty}(\theta-\theta_0)\,\xi(\theta|\underset{\sim}{x})\,d\theta\ =E(\theta|\underset{\sim}{x})-\theta_0$$

故に、

(3.1)　　$E(\theta|\underset{\sim}{x})\leqq\theta_0$　の時、　d_0を選択

　　　　　　　　　$>\theta_0$　の時、　d_1を選択

のように、とるべき決定を決める。

§4. 両側検定。

この節では、$\theta_1\leqq\theta_2$なる２つの実数θ_1とθ_2に対して、２つの仮説を $H_0:\theta_1\leqq\theta\leqq\theta_2$　対 $H_1:\theta<\theta_1$ 或いは、$\theta_2<\theta$ とした時の決定問題を考える。$\Omega_0=[\theta_1,\theta_2]$とする。ここでは、積距離損失関数(R. S., Singh & Wei Laisheng[2]を参照)を仮定する。すなわち、aを、a>0なる定数として、

$$l(d_0,\theta)=\begin{cases}a(\theta-\theta_1)(\theta-\theta_2), & \text{if } \theta\notin\Omega_0\\ 0, & \text{if } \theta\in\Omega_0\end{cases}$$

$$l(d_1,\theta)=\begin{cases}0, & \text{if } \theta\notin\Omega_0\\ a(\theta-\theta_1)(\theta_2-\theta), & \text{if } \theta\in\Omega_0\end{cases}$$

とする。事後確率（密度）関数$\xi(\theta|\underset{\sim}{x})$を用いて、条件付期待損失は、それぞれ、

$$R(d_0|\underset{\sim}{x})=a\{\int_{-\infty}^{\theta_1}(\theta-\theta_1)(\theta-\theta_2)\,\xi(\theta|\underset{\sim}{x})\,d\theta+\int_{\theta_2}^{\infty}(\theta-\theta_1)(\theta-\theta_2)\xi(\theta|\underset{\sim}{x})\,d\theta\}$$

$$R(d_1|\underset{\sim}{x})=a\int_{\theta_1}^{\theta_2}(\theta-\theta_1)(\theta-\theta_2)\xi(\theta|\underset{\sim}{x})\,d\theta$$

と書けるので、

$$\begin{aligned}a^{-1}\{R(d_0|\underset{\sim}{x})-R(d_1|\underset{\sim}{x})\}&=E[(\theta-\theta_1)(\theta-\theta_2)|\underset{\sim}{X}=\underset{\sim}{x}]\\&=E[(\theta-E(\theta|\underset{\sim}{x}))^2]+(E(\theta|\underset{\sim}{x})-\theta_1)(E(\theta|\underset{\sim}{x})-\theta_2)\\&=V(\theta|\underset{\sim}{x})+(E(\theta|\underset{\sim}{x})-\theta_1)(E(\theta|\underset{\sim}{x})-\theta_2)\end{aligned}$$

故に

(4.1)　　$V(\theta|\underset{\sim}{x})+(E(\theta|\underset{\sim}{x})-\theta_1)(E(\theta|\underset{\sim}{x})-\theta_2) \leqq 0$ の時、d_0を選択

>0 の時、d_1を選択

のように、とるべき決定を決める。

§5. 例題。

この節では、§3と、§4についての２つの例題を考える。例１では、$f(x|\theta)$と、$\xi(\theta)$が、共に、正規分布をする場合、例２では、$f(x|\theta)$が、ベルヌイ分布をし、$\xi(\theta)$が、ベータ分布をする場合を考える。これらの$\xi(\theta)$は、いづれも、自然共役分布族と言われるものである。

例１）　$\theta=\theta$が与えられた時、$X_1, \ldots, X_n$を、平均θ、分散１の正規分布に従うランダム標本とし、θの事前分布は、同じく、平均 μ_0（μ_0は、既知）で、分散４の正規分布をする場合を考える。この時、事後密度関数$\xi(\theta|\underset{\sim}{x})$は、

$$\xi(\theta|\underset{\sim}{x}) \propto f(x_1, \ldots, x_n|\theta)\xi(\theta) \propto \exp[-(\theta-\mu_1)^2/(2\sigma_1^2)]$$

で与えられる。　ここで、

$$\mu_1=(4n\bar{x}+\mu_0)/(4n+1)、\qquad \sigma_1^2=4/(4n+1)$$

とする。(但し、$\bar{x}=n^{-1}\sum_{i=1}^{n}x_i$。)

(a)　2つの仮説を、$H_0:\theta\leqq\theta_0$ 対 $H_1:\theta>\theta_0$（θ_0は、実定数）とした時の決定問題を考える。損失関数を、§3のような線型損失関数を考えると、(3.1)より、

$\mu_1 \leqq\theta_0$の時、d_0を選択

$>\theta_0$の時、d_1を選択

のようにとるべき決定を決める。

(b)　２つの仮説を、$H_0:\theta_1\leqq\theta\leqq\theta_2$ 対 $H_1:\theta<\theta_1$ 或いは、$\theta_2<\theta$（θ_1とθ_2は、$\theta_1\leqq\theta_2$なる実数）とした時の決定問題を考える。§4のような積距離損失関数を考えると、(4.1)より、

$\sigma_1^2+(\mu_1-\theta_1)(\mu_1-\theta_2) \leqq 0$の時、$d_0$を選択

>0の時、d_1を選択

のようにとるべき決定を決める。

例２）　$\theta=\theta$が与えられた時、$X_1, \ldots, X_n$ は、独立で、同じベルヌイ分布

$$f(x|\theta)=\theta^x(1-\theta)^{1-x} \quad \text{for } x=0,1$$

$(0<\theta<1)$に従う確率変数とする。パラメータθの事前分布は、自然共役分布族のベータ分布 $\mathrm{Beta}(\alpha_{01}, \alpha_{02})$ $(\alpha_{0i}>0,\ i=1,2)$

$$\xi(\theta)=(1/B(\alpha_{01}, \alpha_{02}))\theta^{\alpha_{01}-1}(1-\theta)^{\alpha_{02}-1} \quad \text{for } 0<\theta<1$$

であるとする。ここで、

$$B(\alpha_{01}, \alpha_{02})=\int_0^1 \theta^{\alpha_{01}-1}(1-\theta)^{\alpha_{02}-1}\,d\theta。$$

この時、同時確率関数は、

$$f(\underset{\sim}{x}|\theta)=f(x_1,\ \dots\ ,x_n|\theta)$$
$$=\theta^{\Sigma_{i=1}^{n}x_i}(1-\theta)^{n-\Sigma_{i=1}^{n}x_i},\quad \text{for } \underset{\sim}{x}\in\{\underset{\sim}{x}:\ x_i=0,1\ (i=1,\ \dots\ ,n)\}$$

$T=\Sigma_{i=1}x_i$ は、充足統計量で、Tの標本分布は、2項分布

$$P(T=t|\theta)=\binom{n}{t}\theta^{t}(1-\theta)^{n-t},\qquad t=0,1,2,\ \dots\ ,n$$

であるので、T=tが与えられた時のθの事後密度関数は、

$$\xi(\theta|t)\propto P(T=t|\theta)\xi(\theta)$$
$$\propto\theta^{\alpha_{11}-1}(1-\theta)^{\alpha_{12}-1}$$

$(\alpha_{11}=\alpha_{01}+t,\ \alpha_{12}=\alpha_{02}+n-t)$となるので、T=tが与えられた時の$\theta$の事後分布は、$Beta(\alpha_{11},\alpha_{12})$である。従って、

$$\mu_2=E(\theta|T=t)=\alpha_{11}/(\alpha_{11}+\alpha_{12}),\quad \sigma_2{}^2=V(\theta|T=t)=\alpha_{11}\alpha_{12}/(\alpha_{11}+\alpha_{12})$$

である。

(a) 2つの仮説$H_0:(0\leq)\theta\leqq\theta_0$ 対 $H_1:\theta<\theta_0(\leq1)$とした時の決定問題を考える。損失関数を§3のような線型損失関数を考えると、(3.1)より、

$\mu_2\ \leqq\theta_0$の時、d_0を選択

$>\theta_0$の時、d_1を選択

のようにとるべき決定を決める。

(b) 2つの仮説を$H_0:\theta_1\leqq\theta\leqq\theta_2$ 対 $H_1:0<\theta\leqq\theta_1$ 或いは、$\theta_2\leqq\theta<1$とした時の決定問題を考える。§4 のような積距離損失関数を考えると、(4.1)より、

$\sigma_2{}^2+(\mu_2-\theta_1)(\mu_2-\theta_2)\ \leqq0$ の時、d_0を選択

>0 の時、d_1を選択

のようにとるべき決定を決める。

参考文献：

[1] Singh, R. S.(1995). Empirical Bayes linear loss hypothesis testing in a nonregular exponential family. Journal of Statistical Planning and Inference, 43, pp. 107-120.

[2] Singh, R. S. & Wei Laisheng(2000). Nonparametric empirical Bayes procedures, asymptotic optimality and rates of convergence for two-tail tests in exponential family., Journal of Nonparametric Statistics, 12, pp. 475-501.

[3] 鈴木雪夫著(1978).「統計解析」 築摩書房