No. 1075

Nonparametric BayesianからみたBayes検定論

by
野上佳子

February 2004

Nonparametric Bayesian からみたBayes検定論

野　上　佳　子

February 5, 2004

1. 考察。

H. Robbinsを、中心とした経験的ベイズ (or nonparametric Bayesianともいう) 或いは、複合決定(Compound Decision) 推論の研究者の方々 (cf. 日本人研究者では、H. 藤本(1976 etc), K. 宮沢(1976 etc), Y. Nogami(1975 etc), S. 森口, G. 鈴木、外国人では、J. Hannan, J. Van Ryzin, V. Susurla, R. S. Singh, D. Gilliland, E. Samuel, M. V. Johns, Jr., T. Obrian etc.) には、私の唱えるBayes推論について、異論を唱える研究者もいるかも知れない。然し乍ら、経験的Bayes推論は、漸近的な最適性を示す理論なのでnが大きくないと、最適性は、示せない。又、収束速度は、n以外の、モデルに関係する値にも依存し、どの位の標本数が、必要なのかについても判然としない点がある。従って、nが大きくない時は、むしろ、験前分布 (先験分布 or 事前分布) $\xi(\theta)$を既知と考えるBayes推論が、好ましく思う。

経験的ベイズ推論の為に仮定する条件は、藤本etc(1976, p.128)を参照されたい。特に、ここで、初めの実験に対する$\tilde{\theta}$の未知の値が、θ_1であるとし、$f(x|\theta_1)$をもつ観測可能な確率変数を、$\tilde{x}_1$で示すことにする。しかし、このθ_1は、未知の験前分布$\xi(\theta)$を持つ確率変数$\tilde{\theta}$の１つの実現値であるが、それは、未知なのであるから、$\xi(\theta)$をもつθの初めの実験に於けるものであることを明示するために、$\tilde{\theta}_1$で表わすことにする。そこで、この実験結果を指示する確率変数の組を、$(\tilde{\theta}_1,\tilde{x}_1)$で表わし、以下同様に、$(\tilde{\theta}_1,\tilde{x}_1)$, ... ,$(\tilde{\theta}_n,\tilde{x}_n)$,$(\tilde{\theta}_{n+1},\tilde{x}_{n+1})$, ...のような実験結果を、指示する確率変数の組の一つの系列が、得られる。ここに、$\tilde{\theta}_1,\tilde{\theta}_2$, ...は、未知の験前分布関数$\xi(\theta)$を持つということが知られるだけで観測可能ではなく、潜在的であって、目的に対しては、観測可能な$(\tilde{x}_1,\ldots,\tilde{x}_n,\ldots)$だけが、利用可能である。(以上、藤本etc.(1976, p.128))

そこで、この験前分布$\xi(\theta)$は、未知であるが、properな分布であることが必要である。というのは、所謂、Bayesian統計学者の承認する無情報験前 (先験)分布 は、nonparametric Bayesian には、その議論の性質上、承認できないものであるということに注目したい。ここで、提唱される新しいBayes推論は、この経験的ベイズアプローチ(nonparametric Bayesian) の考えの中から生まれたものであることを、明記しておきたい。

従って、non-Bayesian (non-empirical Bayesianも含めて) については、同じサイコロを何回も投げる実験を考えると、そのサイコロの１つの目がでる確率、例えば、Pr[目が１] $=\theta_1$とすると、このθ_1上に確率１をもつ$\xi(\theta_1)=1$を想定すればよく、矛盾がないことになる。このように考えると、ここで考える新しいBayes推論 (主に 、検定) は、non-Bayesの理論を包括することに注目されたい。

次の節で回帰モデルへの１つの例を述べておく。（他にも、Y. Nogami(2000)を参照されたい。）

§2. 回帰モデル。

観測モデルを、$\tilde{y}_i=\alpha+\beta x_i+\tilde{\varepsilon}_i$, $i=1,2,\ldots,n$, かつ、$\tilde{\varepsilon}_i \sim_{iid}$ 正規分布$N(0,\sigma^2)$, (σ^2:既知) $i=1,\ldots,n$ とする。（i.e. $\tilde{y}_i \sim_{i.d.} N(\alpha+\beta x_i,\sigma^2)$, $i=1,2,\ldots,n$) 今、$\hat{y}_i=\hat{\alpha}+\hat{\beta}x_i$, $i=1,\ldots,n$とすると、$\hat{\alpha}$と$\hat{\beta}$は、残差平方和$\Sigma_{i=1}^{n}(\tilde{y}_i-\hat{y}_i)^2$を最小にするように選ばれる。（最小二乗推定量）すなわち、

$$\hat{\alpha}=\bar{y}-\hat{\beta}\bar{x} \qquad (\text{or } \tilde{\hat{\alpha}}=\bar{\tilde{y}}-\tilde{\hat{\beta}}\bar{x})$$

$$\hat{\beta}=\Sigma_{i=1}^{n}(y_i-\bar{y})(x_i-\bar{x})/\Sigma_{i=1}^{n}(x_i-\bar{x})^2 \quad (\text{or } \tilde{\hat{\beta}}=\Sigma_{i=1}^{n}(\tilde{y}_i-\bar{\tilde{y}})(x_i-\bar{x})/\Sigma_{i=1}^{n}(x_i-\bar{x})^2)$$

ここで、$\bar{\tilde{y}}=n^{-1}\Sigma_{i=1}^{n}\tilde{y}_i$とする。回帰モデルの性質より、

(i) $E(\tilde{\hat{\beta}})=\beta$ かつ、$\tilde{\hat{\beta}} \sim N(\beta,\ \sigma^2/\Sigma(x_i-\bar{x})^2)$,

(ii) $E(\tilde{\hat{\alpha}})=\alpha$ かつ、$\tilde{\hat{\alpha}} \sim N(\alpha,\sigma^2\Sigma_{i=1}^{n}x_i^2/\{n\Sigma_{i=1}^{n}(x_i-\bar{x})^2\})$

etc.がえられる。(i)についてのみ考える。

$\tilde{\beta}$が、験前分布 $N(\mu_0,\sigma_0^2)$に従うとしよう。そして、仮説$H_0:|\beta-\beta_0|\leqq 0.01$対、対立仮説$H_1:|\beta-\beta_0|>0.01$を考える。この時の損失関数$l(d,\beta)$は、E. Samuel(1963)より、$d_0$を、$H_0$を採択する決定、$d_1$を、$H_1$を採択する決定とし、$b=0.01$とおくと、

$$l(d_0,\beta)=\begin{cases} 0, & \text{if } |\beta-\beta_0|\leqq b, \\ \{(\beta-\beta_0)^2-b^2\}, & \text{if } |\beta-\beta_0|>b \end{cases}$$

$$l(d_1,\beta)=\begin{cases} \{(\beta-\beta_0)^2-b^2\}, & \text{if } |\beta-\beta_0|\leqq b, \\ 0, & \text{if } |\beta-\beta_0|>b. \end{cases}$$

従って、事後確率密度関数は、

$$\xi(\beta|\hat{\beta}) \propto \exp\{-(\beta-\mu_1)^2/\{2\sigma_1^2\}\},$$

但し、

$$\sigma_1^2=\{\Sigma(x_i-\bar{x})^2/\sigma^2+1/\sigma_0^2\}^{-1},$$

$$\mu_1=\{\sigma_0{}^2\Sigma_i(x_i-\bar{x})^2\hat{\beta}+\mu_0\sigma^2\}/\{\sigma_0{}^2\Sigma(x_i-\bar{x})^2+\sigma^2\}.$$

ここで、d_1の事後期待損失を、

$$R(d_1|\hat{\beta})=\int_{\beta_0-b}^{\beta_0+b}\{b^2-(\beta-\beta_0)^2\}\ \xi(\beta|\hat{\beta})\ d\beta=-I_1,$$

とおくと、d_0のそれは、

$$R(d_0|\hat{\beta})=E\{(\tilde{\beta}-\beta_0)^2|\hat{\beta}\}\ -b^2\ -I_1,$$

$$=E\{(\hat{\beta}-E(\tilde{\beta}|\hat{\beta}))^2|\hat{\beta}\}+(E(\tilde{\beta}|\hat{\beta})-\beta_0)^2-b^2-I_1,$$

$$=\sigma_1{}^2+(\mu_1-\beta_0)^2-b^2-I_1$$

となる。 ここで、

$$I_1=\int_{\beta_0-b}^{\beta_0+b}(\beta-(\beta_0+b))(\beta-(\beta_0-b))\ \xi(\beta|\hat{\beta})\ d\beta.$$

我々の検定は、もし、

$R(d_0|\hat{\beta})\leqq R(d_1|\hat{\beta})$ ならば、d_0を選択、又、

$R(d_1|\hat{\beta})>R(d_0|\hat{\beta})$ならば、$d_1$を選択。

$\Phi(z)$は、標準正規分布の累積分布関数で、$\phi(z)$は、その密度関数とし、$z=(\beta-\mu_1)/\sigma_1$, $z_1=(\beta_0-b-\mu_1)/\sigma_1$ かつ、$z_2=(\beta_0+b-\mu_1)/\sigma_1$とおくと、$I_1$は、

$$I_1=\sigma_1{}^2[1+z_1z_2(\Phi(z_2)-\Phi(z_1))+z_1\phi(z_2)-z_2\phi(z_1)]$$

をうる。これより、$R(d_0|\hat{\beta})$と、$R(d_1|\hat{\beta})$が、計算できる。

参考文献

藤本, H. etc. (1976). 決定の数理。筑摩書房。

宮沢, K. (1976). 情報・決定理論序説。岩波書店。

Nogami, Y. (1975). A nonregular squared-error loss set-compound estimation problem., RM-345, Department of Statistics and Probability, Michigan State University. (Ph. D. Dissertation)

野上, Y. (2000). Bayes検定についての考察－決定理論的アプローチー。Discussion Paper Series No. 968, Inst. of Policy and Planning Sciences, Univ. of Tsukuba, January.

Samuel, E. (1963). An empirical Bayes approach to the testing of certain parametric hypotheses., Ann. Math. Stat., Vol. 34