数理解析研究所講究録
590 巻 1986 年 98-107

(特性関数の推定にもとづく経験的ベイズ NON-REGULAR 問題)

AN EMPIRICAL BAYES SQUARED-ERROR LOSS ESTIMATION PROBLEM IN NONREGULAR FAMILIES OF DISTRIBUTIONS.

by Yoshiko Nogami (野上佳子)

University of Tsukuba

1. Introduction.

An initial discussion of the empirical Bayes (EB) problem is given by Robbins (1955, 1963, 1964). Johns and Van Ryzin(1972), Singh(1974, 1976) and Nogami(1983, 1984, 1985) among others considered EB solutions involving kernel-type (see Parzen(1962)) density estimation. On the other hand, O'Bryan and Susarla(1975) and Susarla and O'Bryan(1975) considered EB solutions depending on the estimation of a marginal density using the inversion formula for some absolutely integrable characteristic functions.

In this paper we consider EB solutions using O'Bryan and Susarla (1975)'s method for the squared-error loss under the uniform distributions $U[\theta,\theta+1)$ and $U(0,\theta)$. We first introduce an estimator of a marginal density (in Section 2) and shall obtain convergence rates $O(n^{-\frac{1}{2}})$ for $U[\theta,\theta+1)$ ($\theta\varepsilon[c,d]$ with $-\infty<c\leq d<+\infty$) (Section 3) and $O(n^{-\frac{1}{2}}\log n)$ for $U(0,\theta)$ ($\theta\varepsilon(0,\infty)$) (Section 4), respectively. We also remark (in Section 5) that the mean squared error (MSE) of the nonparametric density estimator under a location parameter family of certain gamma distributions has an upper bound $O(n^{-1+(2\alpha-1)^{-1}})$ for $\alpha>1$.

In the empirical Bayes estimation problem, there is a sequence $\{(X_j,\Theta_j)\}$ of independent random vectors where the unobservable Θ_j are iid according to an unknown prior G and, X_j is independently distributed according to density p_{θ_j} conditional on $\Theta_j=\theta_j$. Let $\doteq$ be a defining property and let $X\doteq X_{n+1}$. By defining a nonrandomized estimator for $\theta\doteq\theta_{n+1}$ by $t_n(X)=t(X_1, X_2, \ldots ,X_n;X)$ in the (n+1)st problem the risk of t_n is given by $R(t_n, G) =\underset{\sim}{E}(t_n(X) -\Theta)^2$ where $\underset{\sim}{E}$ denotes expectation with respect to (wrt) all random variables $\{(X_j,\Theta_j)\}$. With $R\doteq R(G)$ denoting the infimum Bayes risk in the identical component problem, when $R(t_n, G)$ and R are both finite, we have

(1.1) $\qquad (0\leq)\ R(t_n,G) - R = \underset{\sim}{E}(\phi_G(X) - t_n(X))^2$

where

(1.2) $\qquad \phi_G(X) = \int\theta p_\theta(X)\, dG(\theta)/ \int p_\theta(X)\, dG(\theta).$

Since we use Singh's Lemma A.2(1974) in forthcoming sections we state it beforehand without proofs.

<u>Lemma 1.1.</u> (Lemma A.2 of Singh (1974)) Let y, z and L be in $(-\infty,\infty)$ with $z \neq 0$ and $L>0$. If Y and Z are two real valued random variables, then for every $\gamma>0$

$$E(|\tfrac{y}{z} - \tfrac{Y}{Z}|\wedge L)^\gamma \leq 2^{\gamma+(\gamma-1)^+}|z|^{-\gamma}\, E|y - Y|^\gamma + (|\tfrac{y}{z}|^\gamma + 2^{-(\gamma-1)^+}L^\gamma)E|z - Z|^\gamma\}$$

where E means the expectation wrt the joint distribution of (Y,Z) and $a^+=a$ if $a>0$; $=0$ if $a\leq 0$.

As notational convensions $\underset{\sim}{E}_x$ and Var_x denote expectation and variance wrt a random vector $(X_1, \ldots ,X_n,(\Theta|x))$ for given X=x, respectively. Let [A] or A itself denote the indicator function of the set A. A distribution function will also be used to denote the associated measure. Let V and Λ denote the supremum and the infimum, respectively.

2. Estimation of a Marginal Density.

Let $\zeta_j(t)$ be the characteristic function corresponding to the marginal distribution of X_j;

(2.1) $$\zeta_j(t) = E(\exp\{itX_j\}).$$

Since $\zeta_1(t)=\zeta_2(t)= \ldots =\zeta(t)$, we do not exhibit the subscript j. Since $\int|\zeta(t)|\, dt <+\infty$, the marginal density of X is given as follows (cf. Loéve (1963, p. 188)):

(2.2) $$p(x) = \int p_\theta(x)\, dG(\theta) = (2\pi)^{-1}\int e^{-itx}\zeta(t)\, dt.$$

According to O'Bryan and Susarla(1975b) we estimate a truncated pdf of X on [-M, M] (for $0<M<+\infty$) defined by

(2.3) $$p_M(x) = (2\pi)^{-1}\int_{-M}^{M} e^{-itx}\zeta(t)\, dt.$$

Since $p_M(x)= E(n^{-1}\Sigma_{j=1}^{n} \int_0^M \exp\{it(X_j-y)\}\, dt)$,

(2.4) $$p_M(x) = (n\pi)^{-1} \Sigma_{j=1}^{n} (X_j-x)^{-1}\sin(M(X_j- x)).$$

is an unbiased estimator for $p_M(x)$ for given x.

In Sections 3 and 4 we shall find upper bounds for $|2\pi p(y)- 2\pi p_M(y)|^2$ and $\mathrm{Var}_x(\hat{p}_M(y))$ for any y to obtain upper bounds for $\sup_y E_x(\hat{p}_M(y) - p(y))^2$ and shall then apply them for the bounds in Lemma 1.1(Singh(1974)) to get convergence rates for (1.1).

3. Uniform Distribution $U[\theta,\theta+1)$.

For $\theta\varepsilon\,\Omega=[c,d]$ with $-\infty<c\leq d<+\infty$, let $p_\theta(y) =[\theta,\theta+1)$. Throughout this section we denote y-1 by y' and assume that for a positive constant C,

(3.1) $$E(p^{-2}(X)) \leq C\ (< +\infty).$$

For this family, the marginal pdf of X is given by p(y)=G(y)-G(y'). Thus, a telescopic series gives

(3.2) $G(y) = \Sigma_{r=0}^{\infty} p(y-r).$

Since $\int \theta\, p_\theta(x)\, dG(\theta) = \int_{x'}^{x} (x- \int_{(\theta-x')}^{1} dt)\, dG(\theta) = xp(x) - \int_0^1 \int_{x'}^{x'+t} dG(\theta)\, dt,$

(1.2) and (3.2) yields that for $p(x)>0$,

(3.3) $\phi_G(x) = x - \psi(x)$

where

(3.4) $\psi(x) = \int_0^1 \Sigma_{r=0}^{\infty} (p(x'+t-r) - p(x'-r))\, dt/p(x).$

We estimate (3.3) by

(3.5) $\hat{\phi}_M(x) = x - \hat{\psi}_M(x)$

where

(3.6) $\hat{\psi}_M(x) = \{0 \vee \int_0^1 \Sigma_{r=0}^{\infty}(\hat{p}_M(x'+t-r) - \hat{p}_M(x'-r))\, dt/\hat{p}_M(x)\} \wedge 1.$

From (1.1), (3.3) and (3.5)

(3.7) $(0\leq)\ R(\hat{\phi}_M, G) - R = E\{E_{\underset{\sim}{x}}\{(\hat{\psi}_M(X) - \psi(X))^2 \wedge 1\}\}.$

Denoting the quotients of $\hat{\psi}_M(x)$ and $\psi(x)$ by Y/Z and y/z, respectively and applying Lemma 1.1(Singh(1974)) gives

(3.8) $E_{\underset{\sim}{x}}((\hat{\psi}_M(x) - \psi(x))^2 \wedge 1) \leq 8\, p^{-2}(x)(A_x + (3/2)E_{\underset{\sim}{x}}(\hat{p}_M(x) - p(x))^2)$

where

(3.9)
$$A_x = E_{\underset{\sim}{x}} \left| \int_0^1 \Sigma_{r=0}^{\infty} (\hat{p}_M(x'+t-r) - \hat{p}_M(x'-r))\, dt - \right.$$
$$\left. - \int_0^1 \Sigma_{r=0}^{\infty}(p(x'+t-r) - p(x'-r))\, dt \right|^2.$$

To bound rhs(3.8) we introduce following two lemmas:

<u>Lemma 3.1.</u> For $0<M<+\infty$ and any y

(3.10) $|2\pi p(y) - \int_{-M}^{M} e^{-ity}\zeta(t)\, dt| \leq 2(\pi/2)^{1/4} M^{-1/2}.$

<u>Proof.</u>) From (2.2) and the fact that $\zeta(t)=((e^{it}-1)/(it))E(e^{i\Theta t})$, and $|e^{-it(y-\theta)}| \leq 1$,

(3.11) $\quad \text{lhs}(3.10) \le 2\int_M^\infty \left|\frac{e^{it}-1}{it}\right| dt.$

Since for t>0 $|(e^{it}-1)/(it)| = (2(1-\cos t))^{1/2}/t = 2|\sin 2^{-1}t|/t$, changing a variable u=t/2, applying Schwartz's inequality and weakening the range of integration leads to

$$\text{rhs}(3.11) = 2\int_{M/2}^\infty \frac{|\sin u|}{u}\, du$$

$$\le 2\{\int_0^\infty \sin^2 u\, du \int_{M/2}^\infty u^{-2}\, du\}^{\frac{1}{2}}.$$

Since $\int_0^\infty \sin^2 u\, du = (2\sqrt{2})^{-1}\pi^{\frac{1}{2}}$ and $\int_{M/2}^\infty u^{-2} du = 2/M$, the proof is done

<u>Lemma 3.2.</u> For any y

(3.12) $\quad \pi^2 \, \text{Var}_x(\hat{p}_M(y)) \le M\pi n^{-1}.$

<u>Proof.)</u> Since

$$\text{lhs}(3.12) \le n^{-1}\int_{-\infty}^\infty (\int_\theta^{\theta+1}(z-y)^{-2}\sin^2(M(z-y))\, dz)\, dG(\theta)$$

$$\le 2Mn^{-1}\int_0^\infty u^{-2}\sin^2 u\, du.$$

The fact that $2\int_0^\infty u^{-2}\sin^2 u\, du = \pi$ leads to the rhs(3.12).

From above Lemmas 3.1 and 3.2 we get the following lemma:

<u>Lemma 3.3.</u> For $0<M<+\infty$,

$$\sup_y \underset{\sim}{E}_x(\hat{p}_M(y) - p(y))^2 \le 8\pi M^{-1} + M(\pi n)^{-1}.$$

<u>Remark.</u> Hence, we obtain that for $M=n^{1/2}$

(3.13) $\quad \underset{\sim}{E}\,(\hat{p}_M(X) - p(X))^2 \le (8\pi + \pi^{-1})n^{-1/2}.$

To get an upper bound for rhs(3.6), we notice from Hölder's inequality that with s=x' + t

$$\underset{\sim}{E}_x\{\Sigma_{r=0}^\infty(\hat{p}_M(s-r) - p(s-r))\}^2 - \Sigma_{r=0}^\infty \underset{\sim}{E}_x(\hat{p}_M(s-r) - p(s-r))^2$$

$$= \Sigma_{r\ne r^*} \underset{\sim}{E}_x\{(\hat{p}_M(s-r) - p(s-r))(\hat{p}_M(s-r^*) - p(s-r^*))\}$$

$$\le \Sigma_{r\ne r^*} \{\underset{\sim}{E}_x(\hat{p}_M(s-r) - p(s-r))^2 \underset{\sim}{E}_x(\hat{p}_M(s-r^*) - p(s-r^*))^2\}^{\frac{1}{2}}.$$

Applying Lemma 3.3 yields that with N=d- c+2

(3.14) $$\underset{\sim}{E}_x\{\Sigma_{r=0}^{\infty}(\hat{p}_M(s-r) - p(s-r))\}^2 \leq N(N-1)\{8\pi M^{-1} + M(\pi n)^{-1}\}.$$

Thus, since by an exchange of order of integrations and by c_r-inequality (Loéve(1963))

$$A_x \leq 2\{\int_0^1 \underset{\sim}{E}_x\{\Sigma_{r=0}^{\infty}(\hat{p}_M(s-r) - p(s-r))\}^2 \, dt + \underset{\sim}{E}_x\{\Sigma_{r=0}^{\infty}(\hat{p}_M(x'-r) - p(x'-r))\}^2\}.$$

Two applications of (3.14) and Lemma 3.3 leads to

(3.15) $$\text{rhs}(3.8) \leq (32N(N-1) + 12)(8\pi M^{-1} + M(\pi n)^{-1})p^{-2}(x).$$

Thus, (3.15) and (3.1) gives the following theorem:

Theorem 1. For the prior G on [c,d] with the assumption (3.1),

$$(0\leq)\ R(\phi_M, G) - R \leq (32N(N-1)+12)C(8\pi M^{-1}+M(\pi n)^{-1}).$$

We remark that according to Theorem 1, with $M=n^{1/2}$

(3.16) $$(0\leq)\ R(\phi_M, G) - R \leq O(n^{-1/2}).$$

4. Uniform Distribution U(0,θ).

For $\theta\epsilon\,\Omega=(0,\infty)$, let $p_\theta(x)=\theta^{-1}(0,\theta)$. For this family, we assume that

(4.1) $$E(\Theta^{-3/2}) \leq +\infty$$

and

(4.2) $$E(p^{-2}(X)) \leq B\ (<+\infty).$$

Let $P(y)=\int P_\theta(y)\, dG(\theta)$. For this family, $P(x)=G(x)+xp(x)$ and (1.2) gives that for $p(x) > 0$,

(4.3) $$\phi_G(x) = x + \psi(x)$$

where

(4.4) $$\psi(x) = (1 - P(x))/p(x).$$

For each n, $F_n(x)=n^{-1}\Sigma_{j=1}^{n}[X_j\leq x]$ and let $a_n(x)$ be a bounded nonnegative function defined on the positive reals such that for fixed $x(>0)$ $a_n(x)\to\infty$ as $n\to\infty$. Estimate $\psi(x)$ by

(4.5) $$\hat{\psi}_M(x) = \frac{1 - F_n(x)}{\hat{p}_M(x)} \wedge a_n(x).$$

From (4.3), estimate $\phi_G(x)$ by

(4.6) $$\hat{\phi}_M(x) = x + \hat{\psi}_M(x).$$

From (1.1), (4.4) and (4.6) it follows that for sufficiently large n,

(4.7) $$(0\leq)\ R(\hat{\phi}_M,G) - R = E\{E_{\tilde{x}}\{(\hat{\psi}_M(X) - \psi(X))^2 \wedge a_n^2(x)\}\}.$$

But, by Lemma 1.1(Singh(1974)),

(4.8) $$E_{\tilde{x}}\{(\hat{\psi}_M(x) - \psi(x))^2 \wedge a_n^2(x)\} \leq 8p^{-2}(x)\{E_{\tilde{x}}(P(x) - F_n(x))^2 + (3/2)a_n^2(x)E_{\tilde{x}}(p(x) - \hat{p}_M(x))^2\}.$$

To get a bound of rhs(4.8) we shall use forthcoming Lemma 4.3. To get Lemma 4.3 we shall introduce two lemmas.

Lemma 4.1.

(4.9) $$|2\pi p(y) - \int_{-M}^{M} e^{-ity}\zeta(t)\,dt| \leq 4(\pi/2)^{\frac{1}{4}}M^{-\frac{1}{2}}E(\Theta^{-3/2}).$$

Proof.) Since $\zeta(t) = E((e^{it\Theta} - 1)/(it\Theta))$, as in the proof of Lemma 3.1

(4.10) $$\text{lhs}(4.9) \leq 2\int_{-\infty}^{\infty}\int_{M}^{\infty} 2(t\theta)^{-1}|\sin(2^{-1}t\theta)|\,dt\,dG(\theta) \leq 4\int_{-\infty}^{\infty}\theta^{-1}(\int_0^{\infty}\sin^2 u\,du\int_{M\theta/2}^{\infty}u^{-2}\,du)^{1/2}\,dG(\theta)$$

which is rhs(4.9).

Lemma 4.2.

(4.11) $$\pi^2\text{Var}(\hat{p}_M(y)) \leq M\pi n^{-1}.$$

Proof.) As in the proof of Lemma 3.2

$$\text{lhs}(4.11) \leq n^{-1}\int_0^\infty\int_0^\theta (z-y)^{-2}\sin^2(M(z-y))\, dz\, dG(\theta).$$

Since $\int_0^\theta (z-y)^{-2}\sin^2(M(z-y))\, dy \leq M\pi$, the proof is done.

We shall use above two lemmas to obtain following Lemma 4.3:

Lemma 4.3. With $b_0 = 16(\pi/2)^{\frac{1}{2}}(E(\Theta^{-3/2}))^2$,

$$\sup_y \underset{\sim}{E}_x(\hat{p}_M(y) - p(y))^2 \leq b_0 M^{-1} + M\pi n^{-1}.$$

Remark. Hence we obtain that for $M=n^{1/2}$

$$\underset{\sim}{E}(\hat{p}_M(X) - p(X))^2 \leq \{16(\pi/2)^{\frac{1}{2}}(E(\Theta^{-3/2}))^2+\pi\}\, n^{-1/2}.$$

To get an upper bound for (4.7) we notice that $\text{lhs}(4.8) \leq 8p^{-2}(x)\{n^{-1} +(3/2)a_n^2(x)(b_0+\pi)n^{-1/2}\}$. Thus, by (4.3) we obtain

Theorem 2. With $a_n^2(x)=\log n$

(4.12) $\quad (0\leq)\ R(\hat{\phi}_M,G) - R \leq 24\, B(b_0+\pi)n^{-1/2}\log n.$

5. Remark.

For $\theta\varepsilon\,\Omega=(-\infty,\infty)$, let $p_\theta(x)=(\Gamma(\alpha))^{-1}(x-\theta)^{\alpha-1}e^{-(x-\theta)}[x\geq\theta]$ (for $\alpha>1$) and Γ represents the gamma function. For this family, the marginal pdf of X is given by

(5.1) $\quad p(x) = (\Gamma(\alpha))^{-1}\int_{-\infty}^x (x-\theta)^{\alpha-1}e^{-(x-\theta)}\, dG(\theta).$

In this section we shall consider nonparametric density estimation of p(x). To find an upper bound of MSE we shall derive two lemmas.

Lemma 5.1. For $0<M<+\infty$,

(5.2) $\quad |2\pi p(x) -\int_{-M}^M e^{-itx}\,\zeta(t)\, dt| \leq M^{1-\alpha}/(\alpha-1).$

Proof.) Since for t<1, $\zeta(t) = (1-it)^{-\alpha}E(e^{-i\Theta t})$, we have

(5.3) $\quad 2^{-1}\text{lhs}(5.2) \leq \int_M^\infty |1+iu|^{-\alpha}\, du$

$$\leq \int_M^\infty u^{-\alpha}\, du = M^{1-\alpha}/(\alpha-1)$$

which is rhs(5.2).

Lemma 5.2. For any x,

(5.4) $$\pi^2 \, \mathrm{Var}_x(\hat{p}_M(x)) \leq Me^{\alpha-1} \, (\alpha-1)^{1-\alpha} \, n^{-1}.$$

Proof.) Since

(5.5) $$n \ \mathrm{lhs}(5.4) \leq \underset{\sim}{E}_x\{(X_1-x)^{-2} \sin^2(M(X_1-x))\}$$

and since $\sup_{t>0} t^{\alpha-1}e^{-t} = (\alpha-1)^{\alpha-1}e^{-(\alpha-1)}$ $(\alpha>1)$,

(5.6) $$\mathrm{rhs}(5.5) \leq e^{1-\alpha}(\alpha-1)^{\alpha-1} \int_{-\infty}^{\infty}\int_{\theta}^{\infty}(y-x)^{-2}\sin^2(M(y-x))\, dy \, dG(\theta).$$

With a change of variable v=M(y-x) and weakening the range of the resulted integral

(5.7) $$\mathrm{rhs}(5.6) \leq Me^{1-\alpha}(\alpha-1)^{\alpha-1} \, 2 \int_0^\infty v^{-2} \sin^2 v \, dv.$$

Since by a complex integration $2 \int_0^\infty v^{-2}\sin^2 v \, dv = \pi$, this gives the asserted bound.

From above Lemmas 5.1 and 5.2 we obtain

(5.8) $$\sup_x \underset{\sim}{E}_x(\hat{p}_M(x) - p(x))^2 \leq (\alpha-1)^{-2}M^{-2(\alpha-1)} + 2e^{\alpha-1}(\alpha-1)^{1-\alpha}M(\pi n)^{-1}.$$

Therefore, we obtain the following thoerem:

Theorem 3. With $M = n^{(2\alpha-1)^{-1}}$ and $\alpha>1$,

(5.9) $$\underset{\sim}{E}(\hat{p}_M(X) - p(X))^2 \leq O(n^{-1+(2\alpha-1)^{-1}}).$$

For this family of distributions, since $Ep^{-2}(X) = \infty$ even if G is a degenerate distribution, we are not able to use Singh(1974)'s Lemma A.2. So we leave for next time to consider EB problem for this family.

REFERENCES

Johns, Jr., M. V. and Van Ryzin, J. (1972). Convergence rates for empirical Bayes two-action problems II. Continuous case. Ann. Math. Statist. 43, 934-947.

Loéve, Michael(1963). Probability Theory (3rd ed.), Van Nostrand, Princeton.

Nogami, Y. (1983). Rate $O((n^{-1}\log n)^{\frac{1}{2}})$ of risk convergence in the empirical Bayes bootstrap estimation: case of retracted distributions. 京都大学・数理解析研究所・講究録507－統計的漸近理論とその応用, 117-134. (Submitted for publication.)

-------- (1984). Convergence rates in the empirical Bayes estimation under the uniform $U(0,\theta)$ and a location parameter family of gamma distributions. 京都大学・数理解析研究所講究録538. -Nonregular Statistical Estimation-, 185-197. (To appear, partially, in Annals of Statistics, 1986)

------- (1985). Exact $n^{-2/3}$ rate in the empirical Bayes estimation: case of retracted distributions. Statistical Theory and Data Analysis, North Holland, 499-518.

O'Bryan, T. and Susarla, V. (1975). An empirical Bayes estimation problem with nonidentical components involving normal distributions. Commun. Statist. 4, 1033-1042.

Parzen, E. (1962). On estimation of a probability density function and mode. Ann. Math. Statist. 33, 1065-1076.

Robbins, H. (1955). An empirical Bayes approach to statistics. Proc. Third Berkeley Symp. Math. Statist. Prob. 1, University of California Press, 157-164.

-------- (1963). The empirical Bayes approach to testing statistical hypotheses. Rev. Inst. Internat. Statist. 34, 195-208.

-------- (1964). The empirical Bayes approach to statistical decision problem. Ann. Math. Statist., 35, 1-20.

Singh, R. S. (1974). Estimation of derivatives of the average of μ-densities and sequence compound estimation in exponential families. Technical Report 318, Michigan State University.

-------- (1976). Empirical Bayes estimation with convergence rates in noncontinuous Lebesgue-exponential families. Ann. Statist. 4, 431-439.

Susarla, V. and O'Bryan, T. (1975). An empirical Bayes two action problem with nonidentical components for a translated exponential distribution. Commun. Statist., 4, 767-775.