No. 167 (82-34)

Exact rate $n^{-2/3}$ in the empirical Bayes estimation: case of retracted distributions I.

by

Yoshiko Nogami

September 1982

Exact rate $n^{-2/3}$ in the empirical Bayes estimation: case of retracted distributions I.

by

Yoshiko Nogami

1. Introduction

This paper is a continuation of Nogami (1982a). We consider the empirical Bayes squared-error loss estimation problem proposed by H. Robbins (1955), under the family of retracted distributions treated in the author's papers (1981, 1982b).

Let ξ be Lebesgue measure and let f be integrable function with $(0 <)\ f(\cdot) \leq 1$. Define $q(\theta) = (\int_{\theta}^{\theta+1} f\, d\xi)^{-1}$. Letting $p_\theta = dP_\theta/d\xi$ we denote by P(f) the family of probability measures given by

$$P(f) = \{P_\theta \text{ with } p_\theta = q(\theta)[\theta,\ \theta+1)f, \quad \forall\theta \in \Omega\}$$

where Ω is a real line and we denote the indicator function of a set A by A itself (or [A]). Our density p_θ is a retraction of f to the range θ through $\theta+1$.

We use the same notational convensions as stated in Introduction of Nogami (1982a).

Let G be unknown prior on Ω. Define $p(x) = \int p_\theta(x)\, dG(\theta)$ and $P(x) = \int P_\theta(x)\, dG(\theta)$. Let $X_1, \ldots, X_n$ be n i.i.d. past observations distributed according to P. Let X denote the (n+1)st observation X_{n+1}. Let E be the product measure on the space of $(X_1, \ldots, X_n, (X, \Theta))$, resulting from P^n and the joint distribution of (Θ, X). Let $\phi_G(X)$ denote the Bayes estimator vs G given by

$$(1.1)\qquad \phi_G(x) = \int_{x'+}^{x} \theta\, q(\theta)\, dG(\theta) \Big/ \int_{x'+}^{x} q(\theta)\, dG(\theta)$$

where $x' = x-1$, the affix + is intended to describe the integration as over $(x', x]$ and hereafter we abbreviate it.

The risk of an empirical Bayes estimator t_n for θ is $R(t_n, \theta) = E((t_n(X) - \theta)^2)$ and the Bayes envelope is $R \doteq R(\phi_G, G) = \inf_{\phi} R(\phi, G)$. When $R(t_n, G)$ and R are finite we have

$$(1.2)\qquad (0 \leq) R(t_n, G) - R = E(t_n(X) - \phi_G(X))^2.$$

In this paper we exhibit two EB estimators $\tilde{\phi}_n$ and ϕ_n^* for θ in Sections 2 and 4 and show in Sections 3 and 5 that under respective assumptions AI and AII below, $\tilde{\phi}_n$ and ϕ_n^* both have a convergence rate $O(n^{-2/3})$ of $R(t_n, G) - R$. Since we use the same lower bound as that for $R_n^* - R$ in Section 5 of Nogami (1982a), for both $R(\tilde{\phi}_n, G) - R$ and $R(\phi_n^*, G) - R$, we omit this part. Section 6 is a summary.

Assumption AI:

(i) The support of G is bounded i.e.

$\Omega = [c, d]$ where $-\infty < c < d < +\infty$.

(ii) f^{-1} satisfies Lipschitz condition, i.e.

$\sup \{(v - u)^{-1} |(f(v))^{-1} - (f(u))^{-1}| : u < v\} \leq M(< +\infty)$.

(iii) Letting Q, the measure with the density q wrt G, we assume that for some $(0 <)\ h < 1$ and a positive constant L,

$$\sup_{s \in [c, d+1]} \ \sup_{\substack{\text{a.e.} \\ y \in (s, s+h)}} |Q'(y)| \leq L(< +\infty).$$

(iv) With Q in (iii) and a positive constant B, $\int \{Q(x) - Q(x')\}^{-1} dx \leq B(< +\infty)$.

Assumption AII:

(i) $(0 <)\ m^{-1} \le f(\cdot) \le 1$ a.e. Lebesgue measure,

(ii) AI(iii) , (iii) AI(iv).

Let E_x and E be the conditional product measure on the space of $(X_1, \ldots, X_n, \Theta|x)$ given $X = x$ and the marginal probability measure of X, respectively. Let $\vee$ and $\wedge$ denote the supremum and the infimum, respectively.

2. EB estimator $\tilde{\phi}_n$.

In this section we construct EB estimator $\tilde{\phi}_n$ for θ from the alternative form of (1.1) of $\phi_G(x)$ using kernel type density estimator for the marginal p.d.f. $p(x)$ of X.

Let Q be the distribution function defined by

$$(2.1)\qquad Q(y) = \int_{-\infty}^{y} q(\theta)\ dG(\theta).$$

Applying Lemma 1.1 of Nogami (1981) for (1.1) we have that

$$(2.2)\qquad \phi_G(x) = x - \psi(x)$$

where

$$(2.3)\qquad \psi(x) = \{\int_0^1 Q(x' + t)\ dt - Q(x')\}/\{Q(x) - Q(x')\}.$$

To get an EB estimator for θ we estimate $\phi_G(x)$ through estimating the distribution function Q. By the definition of $p(y)$ we get

$$(2.4)\qquad p(y) = f(y)\cdot\{Q(y) - Q(y')\}$$

and hence by a telescopic series,

$$(2.5)\qquad Q(y) = \sum_{r=0}^{\infty} \frac{p(y-r)}{f(y-r)}$$

To estimate $Q(y)$ we estimate $p(y)$. To do so we introduce a kernel function K which is a real valued function vanishing outside the interval $(0, 1)$. To get a bound for $R(\tilde{\phi}_n, G) - R$ we furthermore assume

$$(2.6) \qquad \int_0^1 K(y)\,dy = 1$$

and $\int_0^1 K^2(y)\,dy < +\infty$. As an example of K we let $K(y) = [0 < y < 1]$. For other examples of K, consult with R. S. Singh (1979a) or E. Parzen (1962). We estimate $p(y)$ by $\hat{p}(y) = (nh)^{-1} \sum_{j=1}^{n} K(h^{-1}(X_i - y))$ where $0 < h = h_n (< 1) \to 0$ as $n \to +\infty$. Hence $Q(y)$ can be estimated by

$$(2.7) \qquad \tilde{Q}(y) = \sum_{r=0}^{\infty} \frac{\hat{p}(y-r)}{f(y-r)} .$$

Finally we estimate θ by

$$(2.8) \qquad \tilde{\phi}_n(x) = x - (0 \vee \tilde{\psi}_n(x)) \wedge 1.$$

where

$$(2.9) \qquad \tilde{\psi}_n(x) = \{\int_0^1 \tilde{Q}(x'+t)dt - \tilde{Q}(x')\}/\{\tilde{Q}(x) - \tilde{Q}(x')\}.$$

If we choose $K(y) = [0 < y < 1]$ in the definition of $\hat{p}(y)$, then above $\tilde{\phi}_n$ becomes the same as $\underset{\sim}{\theta}_T$ in Nogami (1975) in the EB problem. Furthermore, if $f \equiv 1$, then $\tilde{\phi}_n$ is the same as ϕ_n^* in Nogami (1982a).

In the next section we shall obtain a rate $O(n^{-2/3})$ for $R(\tilde{\phi}_n, G) - R$.

3. A rate $O(n^{-2/3})$ for $R(\tilde{\phi}_n, G) - R$.

Throughout this section we assume the restriction AI in Section 1. Notice that the assumption AI(ii) is equivalent to the inequality

(3.1) $\left|\frac{f(s)}{f(t)} - 1\right| \leq M\ |s-t|$ for any real s and t.

As usual we use Lemma 4.1 of Singh (1979 b) to obtain a rate $O(n^{-2/3})$ for $R(\tilde{\phi}_n, G) - R$ with a choice of $h = n^{-1/3}$.

In (2.3) and (2.8), let $\psi(x) = v/w$ and $\tilde{\psi}_n(x) = V/W$. In view of (1.2) with t_n replaced by $\tilde{\phi}_n$, and from Lemma 4.1 of Singh (1979b)

$$(3.2)\qquad 0 \leq R(\tilde{\phi}_n, G) - R = E\{E_X(|\frac{V}{W} - \frac{v}{w}| \wedge 1)^2\}$$

$$\leq 8\ E\{|w|^{-2}E_X|v - V|^2\} + 12E\{|w|^{-2}E_X|w - W|^2\}.$$

In view of (2.3) and (2.8), applying the c_r-inequality (Loève (1963, p.155)) and changing the integrations leads to

$$(3.3)\qquad 2^{-1}\ E_x|v - V|^2 \leq \int_0^1 E_x\{Q(x'+t) - Q(x'+t)\}^2\ dt$$

$$+ E_x\{Q(x') - Q(x')\}^2$$

and

$$(3.4)\qquad 2^{-1}\ E_x|w - W|^2 \leq E_x\{\tilde{Q}(x) - Q(x)\}^2 + E_x\{\tilde{Q}(x') - Q(x')\}^2.$$

Letting $\hat{u}(y) = \hat{p}(y)/f(y)$ and $u(y) = p(y)/f(y)$ and applying c_r-inequality after substituting $\pm\ E_x(\hat{u}(x'+t-r))$ we obtain that for any $0 \leq t \leq 1$,

$$(3.5)\qquad 2^{-1}\ E_x\{\tilde{Q}(x'+t) - Q(x'+t)\}^2$$

$$\leq E_x\{\sum_{r=0}^{\infty} (\hat{u}(x'+t-r) - E_x u(x'+t-r))\}^2$$

$$+\{\sum_{r=0}^{\infty} E_x\ \hat{u}(x'+t-r) - u(x'+t-r)\}^2$$

In the forthcoming Lemmas 3.1 and 3.2 we shall get upper bounds for

$E\{u^{-2}(x)$ (the first term of rhs(3.5))} and $E\{u^{-2}(x)$ (the second term of rhs(3.5))} which will in turn give a bound of rhs(3.2) through (3.3) and (3.4).

Lemma 3.1. For any $0 \le t \le 1$ with $A = X' + t$,

(3.6) $$E\{u^{-2}(X) \text{ (the first term of rhs(3.5))}\}$$

$$\le (nh)^{-1}(1+M)(1+M+MN)BQ(\infty)\int_0^1 K^2(y)dy$$

where N is the greatest integer less than d+2-c and M and B are as defined in (3.1) and AI (iv), respectively.

Proof. Expanding the square of the summation in r in the first term of rhs(3.5) leads to

(3.7) $$E_x\{\sum_{r=0}^{\infty} (\hat{u}(s-r) - E_x\hat{u}(s-r))\}^2$$

$$= \sum_{r=0}^{\infty} \mathrm{Var}_x(\hat{u}(s-r))$$

$$+ \sum_{r\neq r^*} \{E_x(\hat{u}(s-r)\hat{u}(s-r^*)) - E_x\hat{u}(s-r)E_x\hat{u}(s-r^*)\}.$$

Changing the variable $y = h^{-1}(x_j-s+r)$ and $z = h^{-1}(x_{j^*}-s+r)$ and noticing, in (3.8), that for $0 \le y \le 1$ and $r^* \neq r$, $K(y + h^{-1}(r^*-r)) = 0$ we obtain

(3.8) $$E_x(\hat{u}(s-r)\hat{u}(s-r^*))$$

$$= (1 - n^{-1})(f(s-r)f(s-r^*)^{-1}\int_0^1\int_0^1 K(y)K(z)$$

$$\cdot\, p(s-r+hy)p(s-r+hz)dydz$$

and

$$(3.9)\qquad E_x\hat{u}(s-r)E_x\hat{u}(s-r^*) = (f(s-r)f(s-r^*))^{-1}\int_0^1 K(y)p(s-r+hy)dy \times \int_0^1 K(z)p(s-r+hz)dz.$$

Thus,

(3.10) the second term of rhs(3.7) ≤ 0.

Similarly,

(3.11) the first term of rhs(3.7)

$$= \sum_{r=0}^{\infty} f^{-2}(s-r)\{(nh)^{-1}\int_0^1 K^2(y)p(s-r+hy)dy - n^{-2}(\int_0^1 K(y)p(s-r+hy)dy)^2\}$$

$$\leq (nh)^{-1}\int_0^1 K^2(y) \sum_{r=0}^{\infty}[\frac{f(s-r+hy)}{f(s-r)} \{Q(s-r+hy) - Q(s'-r+hy)\}/f(s-r)]dy.$$

Therefore, since

$$\{Q(s-r+hy) - Q(s'-r+hy)\}/f(s-r)$$

$$= \frac{Q(s-r+hy)}{f(s-r)} - \frac{Q(s'-r+hy)}{f(s'-r)} + \{\frac{1}{f(s'-r)} - \frac{1}{f(s-r)}\}Q(s'-r+hy)$$

$$\leq \frac{Q(s-r+hy)}{f(s-r)} - \frac{Q(s'-r+hy)}{f(s'-r)} + M\,Q(\infty)$$

where the inequality follows from (3.1), (3.1) and this gives

$$(3.12)\qquad (3.7) \leq (nh)^{-1}(1+Mh)\{\int_0^1 K^2(y)(Q(s+hy)/f(s))dy \quad Q(\infty)\int_0^1 K^2(y)dy\}$$

But, from $f \leq 1$ and AI(iv),

$$(3.13) \qquad Eu^{-2}(x) \leq \int \{Q(x) - Q(x')\}^{-1}\, dx \leq B.$$

Thus, (3.12) with noticing from (3.1) that $f(x)/f(s) \leq 1 + M(x-s)$, weakening the resulted bound, and (3.13) given the rhs(3.6).□

Lemma 3.2. For any $0 \leq t \leq 1$ with $s = X' + t$,

$$(3.14) \qquad E\{u^{-2}(X) \text{ (the second term of rhs(3.5))}\}$$

$$\leq h^2\, B\, (L + 3M\, Q(\infty))^2 \int_0^1 K^2(y)dy$$

where L is defined in AI(ii).

Proof. Since by changing the variable and (2.6)

$$(3.15) \qquad E_x\hat{u}(s-r) - u(s-r) = (f(s-r))^{-1} \int_0^1 K(y)\{p(s-r+hy) - p(s-r)\}dy,$$

we have that

$$(3.16) \qquad \left|\sum_{r=0}^{\infty} (E_x\hat{u}(s-r) - u(s-r))\right|$$

$$= \left|\int_0^1 K(y)\left\{\sum_{r=0}^{\infty} \frac{f(s-r+hy)}{f(s-r)}(Q(s-r+hy) - Q(s'-r+hy)) - \sum_{r=0}^{\infty}(Q(s-r) - Q(s'-r))\right\}\cdot dy\right|$$

$$\leq \int_0^1 |K(y)|\Big[\frac{f(s+hy)}{f(s)}\{Q(s+hy) - Q(s)\}$$

$$+\left|\frac{f(s+hy)}{f(s)} - 1\right|Q(s) - (1 - Mhy)Q(s'+hy)$$

$$+ (1 + Mhy) \sum_{r=1}^{\infty}\{Q(s-r+hy) - Q(s'-r+hy)\}\Big]dy$$

$$\leq \int_0^1 |K(y)|\,[\{Q(s+h) - Q(s)\} + Mhy\{Q(s+hy) + 2Q(s'+hy)\}]dy$$

where each inequality follows from a telescopic series and two applications of (3.1).

Since Q is of bounded variation on $[s, s+\eta]$ (cf. Royden (1968)) with $\eta = hy$, there exists the first derivative Q' a.e. in $(s, s+\eta)$. Furthermore, Q(z) is right continuous. These facts together satisfy the requirements for the following Taylor expansion (cf. Singh (1978, p.639)):

$$(3.17)\qquad Q(s+hy) - Q(s) = \int_s^{s+hy} Q'(x)dz.$$

Thus, using this and AI(iii) and weakening the resulted bound we have

$$(3.18)\qquad \text{rhs}(3.16) \le \int_0^1 |K(y)| \{\int_s^{s+hy} |Q'(x)|dz + 3MhyQ(\infty)\}dy$$

$$\le h\int_0^1 |K(y)|dy\ \{L + 3M\ Q(\infty)\}$$

and therefore

$$(3.19)\qquad (\text{rhs}(3.16))^2 \le h^2 \int_0^1 K^2(y)dy\ \{L + 3M\ Q(\infty)\}^2.$$

This inequality and (3.13) leads to the asserted bound.□

Therefore, (3.2) through (3.5) together with Lemmas 3.1 and 3.2 gives that

$$(3.20)\qquad R(\tilde{\phi}_n, G) - R \le 80\ B \int_0^1 K^2(y)dy\{(nh)^{-1}(1+M)(L+M+MN)Q(\infty)$$
$$+ h^2(L + 3MQ(\infty))^2\}.$$

Thus, we obtain the following theorem:

Theorem 3.1. With the assumption AI and a choice of $h = n^{-1/3}$,

$$R(\tilde{\phi}_n, G) - R = O(n^{-2/3}).$$

In the next section we introduce another EB estimator ϕ_n^* and in Section 5 we show that ϕ_n^* has the same convergence rate $O(n^{-2/3})$ as $\tilde{\phi}_n$, with merely the assumption of boundedness below of f rather than that of bounded Ω and Lipschitz condition for f^{-1}.

4. EB estimator ϕ_n^*

In this section we construct another EB estimator ϕ_n^* whose construction is similar to $\tilde{\phi}_n$ in Section 2.

We estimate (2.2) by estimating Q of form (2.5). Let K be a kernel function given in Section 2. In view of (2.5) we estimate Q by

$$(4.1) \qquad Q^*(y) = \sum_{r=0}^{\infty} (nh)^{-1} \sum_{j=1}^{n} \frac{K(h^{-1}(X_j - y + r))}{f(X_j)}$$

and hence, given X = x, estimate θ by

$$(4.2) \qquad \phi_n^*(x) = x - (0 \vee \psi_n^*(x)) \wedge 1$$

where

$$\psi_n(x) = \{\int_0^1 Q^*(x'+t)dt - Q^*(x')\}/\{Q^*(x) - Q^*(x')\}.$$

As we can see easily, if we choose $K(y) = [0 < y < 1]$, then above ϕ_n^* becomes the same as the ϕ_{nn}^* in Nogami (1981) in the EB problem. Furthermore, if $f \equiv 1$, then such $\phi_n^* = \tilde{\phi}_n$ and is the same as ϕ_n^* in Nogami (1982 a).

5. A Rate $O(n^{-2/3})$ for $R(\phi_n^*, G) - R$

Throughout this section we assume the assumption AII in Section 1.

We proceed in the same way as in Section 2 until (3.5) with the superscript ∿ replaced by *, but $\hat{u}(y)$ is now form $\hat{u}(y) = (nh)^{-1} \sum_{j=1}^{n} \frac{K(h^{-1}(X_j - y))}{f(X_j)}$. In the similar way that we have done in Lemmas 3.1 and 3.2, we obtain Lemmas 5.1 and 5.2.

Lemma 5.1 For any $0 \le t \le 1$ with $s = X'+t$,

$$(5.1) \qquad E\{u^{-2}(x) \text{ (the first term of rhs(3.5))}\} \le (nh)^{-1} m\ B\ Q(\infty) \int_0^1 K^2(y)dy$$

where m and B are given in the assumption AII.

Proof. By the similar methods to obtain (3.10) and (3.11) we can get that

$$(5.2)\qquad (nh)(3.7) \le \sum_{r=0}^{\infty} h^{-1}\int_{s-r}^{s-r+h} K^2(h^{-1}(z-s+r))\{Q(z) - Q(z')\}/f(z)dz.$$

By AII(i), a change of a variable $y = h^{-1}(z-s+r)$, and a telescopic series,

$$(5.3)\qquad \text{rhs}(5.2) \le m\int_0^1 K^2(y) \sum_{r=0}^{\infty} \{Q(s-r+hy) - Q(s'-r+hy)\}dy$$

$$= m\int_0^1 K^2(y)Q(s+hy)dy.$$

Thus, weakening the above bound and using (3.13) gives us Lemma 5.1.□

Lemma 5.2. For any $0 \le t \le 1$ with $s = X' + t$,

$$(5.4)\qquad E\{u^{-2}(x)(\text{the second term) of rhs}(3.5))\}$$

$$\le h^2 B\ L^2 \int_0^1 K^2(y)dy.$$

Proof. In the same way as (3.16) we have that

$$(5.5)\qquad \left|\sum_{r=0}^{\infty} (E_x\hat{u}(s-r) - u(s-r))\right| = \int_0^1 K(y)\Big[\sum_{r=0}^{\infty} \{Q(s-r+hy) - Q(s'-r+hy)\}$$

$$+ \sum_{r=0}^{\infty} \{Q(s-r) - Q(s'-r)\}\Big]\ dy.$$

By two telescopic series, (3.17) and applying AII(ii),

$$(5.5) = \int_0^1 K(y)\{Q(s+hy) - Q(s)\}dy$$

$$\le \int_0^1 K(y) \int_s^{s+hy} |Q'(z)|dz\ dy$$

$$\le L\ h \int_0^1 K(y)dy.$$

Therefore, from $(\int_0^1 K(y)dy)^2 \leq \int_0^1 K^2(y)dy$, $f \leq 1$ and (3.13), we obtain the asserted bound.□

Therefore, (3.2) through (3.5) together with Lemmas 5.1 and 5.2 gives us that

$$\text{(5.6)} \qquad R(\phi_n^*, G) - R \leq 80\ B \int_0^1 K^2(y)dy\{(nh)^{-1}mQ(\infty) + h^2L^2\}.$$

Thus, we obtain the same convergence rate as $\tilde{\phi}_n$ as follows:

Theorem 5.1. With the assumption AII and a choice of $h = n^{-1/3}$,

$$R(\phi^*_n, G) - R = O(n^{-2/3}).$$

Remark. As we can see from (5.2), the assumption AII(i) can be replaced by (i)' $\sup_v \sup_{\substack{a.e. \\ y \in (v,v+h)}} (f(y))^{-1} \leq S(< +\infty)$ which is weaker than AII(i). In this case, B in rhs(5.6) is replaced by S. Although $\tilde{\phi}_n$ and ϕ_n^* have the same rate $O(n^{-2/3})$ of convergence, when we compare two bounds in (3.20) and (5.6) (with B replaced by S) we can see that the bound of (5.6) is better than that of (3.20). However, the structure of the basic estimate $\tilde{Q}(y)$ for $Q(y)$ may be more natural than $Q^*(y)$.

6. Summary

In case of application we will be interested in the form of the function f. Actually, restrictions AI(ii) and AII(i) are strong enough to limit the function f. For example, among others f could be a triangular p.d.f. or a trigonometric function, but cannot be normal p.d.f. or gamma p.d.f. which will often occur in practrical problems. So we need to consider these cases separately from the result here.

References

1. Loève, M. (1963). Probability Theory (3rd ed.), Van Nostrand, Princeton.

2. Nogami, Y. (1975). A non-regular squared-error loss set compound estimation problem, RM-345, Department of Statistics and Probability, Michigan State University.

3. ________ (1981). The set-compound one-stage estimation in the nonregular family of distributions over the interval $[\theta, \theta+1)$, Ann. Inst. Statist. Math., 33, A, 67-80.

4. ________ (1982a). Exact rate $n^{-2/3}$ in the empirical Bayes estimation: case of uniform distributions I, D.P. 150 (82-17), Inst. of Socio-Economic Planning, Univ. of Tsukuba.

5. ________ (1982b). A rate of convergence for the set compound estimation in a family of certain retracted distributions, Ann. Inst. Statist. Math., 34, No.2, A, 241-257.

6. Parzen, Emanuel (1962). On estimation of a probability density function and mode. A. M. S., 33, 1065-1076.

7. Robbins, H. (1975). An empirical Bayes approach to statistics. Proc. Third Berkeley Symp. Math. Statist. Probability, 1, 157-163.

8. Royden, H. L. (1968). Real Analysis (2nd ed.). Macmillan Company, New York.

9. Singh, R. S. (1978). Nonparametric estimation of derivatives of average of μ-densities with convergence rates and applications. SIAM J. Appl. Math., 35, 637-649.

10. ________ (1979a). Mean squared errors of estimators of a density and its derivatives, Biormetrica, 66, No. 1, 177-180.

11. ________ (1979b). Empirical Bayes estimation in Lebesgue-exponential families with rates near the best possible rate. A.S., 7, No. 4, 890-902.