Ann. Inst. Statist. Math.
34 (1982), Part A, 241-257

A RATE OF CONVERGENCE FOR THE SET COMPOUND ESTIMATION IN A FAMILY OF CERTAIN RETRACTED DISTRIBUTIONS*

YOSHIKO NOGAMI

(Received Dec. 17, 1979; revised Jan. 11, 1982)

Summary

This paper is concerned with the set compound squared-error loss estimation problem. Here, the author obtains Lévy consistent estimate $\hat{G}_n$ of the empiric distribution G_n of the parameters $\theta_1, \cdots, \theta_n$ for a more general family of retracted distributions on the interval $[\theta, \theta+1)$ than the uniform on $[\theta, \theta+1)$ as in R. Fox (1970, *Ann. Math. Statist.*, **41**, 1845–1852; 1978, *Ann. Statist.*, **6**, 846–853) and exhibits a decision procedure based on $\hat{G}_n$ with a convergence rate $O((n^{-1}\log n)^{1/4})$ for the modified regret uniformly in $(\theta_1, \theta_2, \cdots, \theta_n) \in \Omega^n$ with bounded Ω. The author also gives a counterexample to the convergence of the modified regret for $\Omega=(-\infty, \infty)$.

0. Introduction

The set compound problem simultaneously considers n statistical decision problems each of which is structually identical to the component problem. The loss is taken to be the average of n component losses.

Let ξ be Lebesgue measure and f an integrable function with $0 \leqq f \leqq 1$. Define $q(\theta)=\left(\int_\theta^{\theta+1} f d\xi\right)^{-1}$ and assume that q is uniformly bounded by a finite constant, say m. Letting $p_\theta = dP_\theta/d\xi$ we denote by $\mathcal{P}(f)$ the family of probability measures given by

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

where Ω is a real interval and we denote the indicator function of a set A by A itself. The family of probability measures as above are useful when the population has the distribution with a Lebesgue density f on $(-\infty, \infty)$ in general (as Normal distribution), but the range of observations drawn as a sample is no more than 1 and furthermore

* This is part of the author's Ph. D. Thesis at Michigan State University.

the right end point of that range is a parameter θ (although our density $p_\theta(x)$ is a retraction of f to the range θ through $\theta+1$).

In this paper, the component problem considered is the squared-error loss estimation of θ based on X with distribution $p_\theta \in \mathcal{P}(f)$. For any prior distribution G on Ω, let $R(G)$ be the Bayes risk versus G in this component problem.

Let $X_1, \cdots, X_n$ be n independent random variables with X_j distributed according to $P_{\theta_j} \in \mathcal{P}(f)$. Let $\boldsymbol{t}=(t_1, t_2, \cdots, t_n)$ be a set compound procedure: for each $j=1, 2, \cdots, n$, $t_j(\boldsymbol{X})$ is an estimator of θ_j based on $\boldsymbol{X}=(X_1, \cdots, X_n)$. Let G_n denote the empiric distribution of $\theta_1, \cdots, \theta_n$ and let

$$D(\boldsymbol{\theta}, \boldsymbol{t})=\int n^{-1} \sum_{j=1}^{n} (t_j(\boldsymbol{x})-\theta_j)^2 d\boldsymbol{P}(\boldsymbol{x})-R(G_n) \tag{2}$$

where $\boldsymbol{P}=P_{\theta_1}\times \cdots \times P_{\theta_n}$. We shall call $D(\boldsymbol{\theta}, \boldsymbol{t})$ the modified regret for a decision procedure $\boldsymbol{t}$.

With squared-error loss, let $\boldsymbol{\theta}_{G_n}$ be the procedure whose component procedures are Bayes against G_n: $\boldsymbol{\theta}_{G_n}(\boldsymbol{X})=(\theta_{1n}, \theta_{2n}, \cdots, \theta_{nn})$ with, for each j,

$$\begin{aligned}\theta_{jn}&=\int \theta p_\theta(X_j)dG_n(\theta)\Big/\int p_\theta(X_j)dG_n(\theta)\\&=\int_{(X_j-1)+}^{X_j} \theta q(\theta)dG_n(\theta)\Big/\int_{(X_j-1)+}^{X_j} qdG_n\end{aligned} \tag{3}$$

where the affix $+$ is intended to describe the integration as over $(X_j-1, X_j]$. Henceforth we delete $+$ in lower limits of $\int$'s.

If $\sup\{|D(\boldsymbol{\theta}, \boldsymbol{t})|: \boldsymbol{\theta} \in \Omega^n\}=O(n^{-\alpha})$, then we will say that $\boldsymbol{t}$ has a rate $O(n^{-\alpha})$.

R. Fox [3] solved empirical Bayes squared error loss estimation (SELE) problem for the uniform distribution $U[\theta, \theta+1)$ by exhibiting, without rates, a (one-stage) decision procedure estimating the Bayes estimator directly. Y. Nogami [8] extends his result to set compound SELE problem for a family $\mathcal{P}(f)$ of retracted distributions on the interval $[\theta, \theta+1)$, demonstrated the existence of a one-stage decision procedure $\boldsymbol{\phi}^*$ with a convergence rate $O(n^{-1/3})$ and showed that $\boldsymbol{\phi}^*$ has the best exact order $n^{-2/3}$ of convergence of the modified regret $D(\boldsymbol{0}, \boldsymbol{\phi}^*)$ at $f\equiv 1$.

Furthermore, for the uniform $U[\theta, \theta+1)$ case where $\theta \in (-\infty, \infty)$ Fox [2] exhibited a distribution-valued Lévy consistent estimate $\hat{G}_n$ of G_n. In empirical Bayes problem where the θ_i are i.i.d. with common distribution G, Fox ([3], the second Remark after Theorem 3.1) indicated (without rates) a convergence of the expected risks to $R(G)$ for a (bootstrap or two-stage) decision procedure $\hat{\theta}$ based on component procedures Bayes versus an estimate $\hat{G}_n$. In this paper we exhibit a distribution-valued Lévy consistent estimate $\hat{G}_n$ of G_n for the family $\mathcal{P}(f)$ of retracted distributions and obtain a convergence rate $O((n^{-1}\log n)^{1/4})$ for the modified regret $D(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$, when Ω is bounded. We also give a counterexample to $D(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})\to 0$ when $\Omega=(-\infty, \infty)$.

In Section 1 (Theorem 1.1) we exhibit an upper bound of the modified regret $D(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$ (uniform wrt $\boldsymbol{\theta}\in\Omega^n$) in terms of Lévy metric $L(G_n, \hat{G}_n)$ of G_n and any distribution-valued estimate $\hat{G}_n$, when Ω is bounded. In Section 2 we construct a particular distribution-valued Lévy consistent estimate $\hat{G}_n$ of G_n for $\Omega=(-\infty, \infty)$. Under an additional assumption that $1/f$ satisfies the Lipschitz condition, we show in Theorem 2.1, by making use of the bound in Theorem 1.1, that the set compound decision procedure $\hat{\boldsymbol{\theta}}$ based on $\hat{G}_n$ has a rate $O((n^{-1}\log n)^{1/4})$. Section 3 shows in Theorem 3.1 that when $\Omega=(-\infty, \infty)$, there is no sequence of estimates of $\boldsymbol{\theta}$ for which $D(\boldsymbol{\theta}, \boldsymbol{t})$ converges to zero.

Notational conventions

$\boldsymbol{P}$ abbreviate $\times_{j=1}^{n} P_{\theta_j}$. A distribution function also represents the corresponding measure. We often let Ph, $P(h)$ or $P(h(\omega))$ denote $\int h(\omega)dP(\omega)$. R denotes the real line. We abbreviate $y-1$ to y'. We denote the indicator function of a set A by $[A]$ or simply A itself. $\vee$ and $\wedge$ denote the supremum and the infimum, respectively.

1. An upper bound of the modified regret

Let $\Omega=[c, d]$, where $-\infty<c\leqq d<+\infty$, throughout this section. Let $\hat{G}_n$ be a distribution-valued random variable which is an estimate of the empiric distribution G_n, obtained from $X_1, \cdots, X_n$. Define $\hat{\boldsymbol{\theta}}=(\hat{\theta}_1, \cdots, \hat{\theta}_n)$ to be the procedure such that, for each j, $\hat{\theta}_j(\boldsymbol{X})$ $(=\hat{\theta}_{jn})$ is of form (3) with G_n replaced by $\hat{G}_n$ (0/0 is understood to be X_j).

The modified regret for a procedure $\boldsymbol{t}$ is of form

$$D(\boldsymbol{\theta}, \boldsymbol{t})=n^{-1}\sum_{j=1}^{n}\{\boldsymbol{P}(t_j(\boldsymbol{X})-\theta_j)^2-\boldsymbol{P}(\theta_{jn}-\theta_j)^2\} \tag{4}$$

because in (2) $R(G_n)=\boldsymbol{P}\left\{n^{-1}\sum_{j=1}^{n}(\theta_{jn}-\theta_j)^2\right\}$.

Lévy distance for two distribution functions F and H of random variables (cf. Feller [1], p. 285) is defined by

$$(5) \qquad L(F, H)=\wedge\{\varepsilon\geqq 0:\ F(y-\varepsilon)-\varepsilon\leqq H(y)\leqq F(y+\varepsilon)+\varepsilon,\ \text{for all } y\in R\}\ .$$

Remark that the infimum in the definition (5) attains (see Appendix of Nogami [7]).

In this section we shall exhibit an upper bound of the modified regret $|D(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})|$. To do so, the main development is Lemma 1.5 in which we show that the average expectation of $|\hat{\theta}_{jn}-\theta_{jn}|$ over the set where $L(G_n, \hat{G}_n)\leqq\varepsilon$ is bounded by at most a constant times ε with $0<\varepsilon<1$. For the proof of Lemma 1.5 we use Lemma 5.1 of R. S. Singh [12], Proposition A of Nogami [7] and Lemma A.3 of Nogami [7]. Hence, we shall state them here beforehand.

LEMMA 1.1 (R. S. Singh [12]). *Let y, z and B be in R with $z\neq 0$ and $B>0$. If Y and Z are real valued random variables, then for every $\gamma>0$*

$$\mathrm{E}\,(|(Y/Z)-(y/z)|\wedge B)^{\gamma}<2^{\gamma+(\gamma-1)^+}|z|^{-\gamma}\{\mathrm{E}\,|y-Y|^{\gamma}+(|y/z|^{\gamma}+2^{-(\gamma-1)^+}B^{\gamma})\,\mathrm{E}\,|z-Z|^{\gamma}\}$$

where a^+ denotes $0\vee a$.

PROPOSITION (Nogami [7], Proposition A). Let $I=(a, b]$ be a finite interval and let F_I be the retraction of a distribution function F into the closed interval $[F(a+), F(b+)]$. Then,

$$L(F_I, G_I)\leqq|(F-G)(a+)|\vee|(F-G)(b+)|\vee L(F, G)$$

where we use $+$ on the line to denote the right limit.

For Lemma 1.2 we need to introduce following definition:

DEFINITION 1.1. With h, a function defined on a real interval I, the modulus of continuity of h is the function given by

$$\alpha(\varepsilon)=\vee\{h(\omega_1)-h(\omega_2):\ \omega_1, \omega_2\in I,\ |\omega_1-\omega_2|<\varepsilon\}$$

for every $\varepsilon>0$.

In Lemma 1.2 below, the natural generalization of the inverse probability integral transformation is used to develop bounds for the same difference of integrals in Lemma 8′ of Oaten ([9], Appendix) without partitioning as in Oaten's proof.

LEMMA 1.2 (Nogami [7], Lemma A.3). *Let I be a finite interval $\{a, b\}$ supporting finite measures μ and ν and let h be measurable on I into a finite interval $[c, d]$. Let F and G be distribution functions in-*

ducing μ *and* ν *with* $F(a-)\vee G(a-)\leqq F(b+)\wedge G(b+)$. *Then* $\left|\int hd(\mu-\nu)\right|$ *has the following family of bounds*

$$\frac{d-c}{2}\{|(F-G)(a-)|+|(F-G)(b+)|\}+\alpha(L(F,G)+)$$
$$\times\{F(b+)\wedge G(b+)-F(a-)\vee G(a-)\}\frac{|d+c|}{2}|\mu I-\nu I|\ .$$

We now go back to find the bound of (4). Since $X_j'<\theta_{jn}\leqq X_j$ by (3) whatever be the distribution G_n, $|(\hat{\theta}_{jn}-\theta_j)^2-(\theta_{jn}-\theta_j)^2|\leqq 2|\hat{\theta}_{jn}-\theta_{jn}|$. Hence, it follows from (4) that

$$2^{-1}|D(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})|\leqq n^{-1}\sum_{j=1}^{n}\boldsymbol{P}|\hat{\theta}_{jn}-\theta_{jn}|\ . \tag{6}$$

For fixed j, since $|\hat{\theta}_{jn}-\theta_{jn}|\leqq 1$, for any $0<\varepsilon<1$,

$$\boldsymbol{P}|\hat{\theta}_{jn}-\theta_{jn}|\leqq\boldsymbol{P}[L(G_n,\hat{G}_n)>\varepsilon]+\boldsymbol{P}(|\hat{\theta}_{jn}-\theta_{jn}|[L(G_n,\hat{G}_n)\leqq\varepsilon])\ . \tag{7}$$

Before dealing with the second term of rhs (7), we introduce two lemmas.

LEMMA 1.3. *For any* $s,t\in R$ *with* $s\leqq t$ *and for any* $\delta\geqq 0$ *and* $\eta\geqq 0$ *with* $\delta+\eta<1$,

$$n^{-1}\sum_{j=1}^{n}A_j\leqq t-s \tag{8}$$

where for each j,

$$A_j=P_{\theta_j}\left\{(G_n(X_j-s)-G_n(X_j-t))[\theta_j+\delta\leqq X_j<\theta_j+1-\eta]\Big/\int_{X_j'+\eta}^{X_j-\delta}qdG_n\right\}\ .$$

PROOF. Since for each j,

$$A_j=\int\left\{q(\theta_j)[\theta_j+\delta\leqq y<\theta_j+1-\eta]\Big/\int_{y'+\eta}^{y-\delta}qdG_n\right\}$$
$$\times f(y)(G_n(y-s)-G_n(y-t))dt \tag{9}$$

and since $[\theta_j+\delta\leqq y<\theta_j+1-\eta]=[y'+\eta<\theta_j\leqq y-\delta]$, the average wrt $j=1,\cdots,n$ of the numerator in the quotient equals to the denominator. Also, since $f\leqq 1$, taking the average wrt j over (9) and interchanging the integral and average operation leads to lhs (8) $\leqq\int(G_n(y-s)-G_n(y-t))dy$.

But, the Fubini Theorem leads to

$$\int(F(y-s)-F(y-t))dy=\int\int_{u+s}^{u+t}dydF(u)=t-s$$

for an arbitrary distribution function F of a random variable and any $s, t \in R$ with $s<t$. This gives us the resulted bound.

LEMMA 1.4. *For all* $s \in R$,

$$(10)\quad n^{-1}\sum_{j=1}^{n} \boldsymbol{P}\left\{|(G_n-\hat{G}_n)(X_j-s)|[L(G_n, \hat{G}_n)\leqq\varepsilon]\Big/\int_{X'_j}^{X_j} q\, dG_n\right\}\leqq(d-c+3)\varepsilon\ .$$

PROOF. For j fixed we let $z=X_j-s$. By the definition of $L(G_n, \hat{G}_n)$ and the fact that the infimum in the definition of Lévy distance is attained,

$$|(G_n-\hat{G}_n)(z)|[L(G_n, \hat{G}_n)\leqq\varepsilon]\leqq\varepsilon+G_n(z+\varepsilon)-G_n(z-\varepsilon)\ .$$

Thus,

$$\begin{aligned}\text{lhs}\,(10)\leqq\varepsilon&\left\{n^{-1}\sum_{j=1}^{n} P_{\theta_j}\left(\int_{X'_j}^{X_j} q\, dG_n\right)^{-1}\right\}\\&+n^{-1}\sum_{j=1}^{n} P_{\theta_j}\left\{(G_n(X_j-s+\varepsilon)-G_n(X_j-s-\varepsilon))\Big/\int_{X'_j}^{X_j} q\, dG_n\right\}\ .\end{aligned}$$

From the proof of Lemma 1.3 we can see that ε^{-1}\{the first term\} is no more than $d-c+1$. Hence, an application of Lemma 1.3 with $(s, t, \eta, \delta)=(s-\varepsilon, s+\varepsilon, 0, 0)$ to the second term of the rhs completes the proof.

We need Lemma 1.1 (R. S. Singh [12]) in the proof of Lemma 1.5 below which will give us an upper bound of the average wrt j of the second term of rhs (7).

LEMMA 1.5. *For* $\varepsilon>0$,

$$n^{-1}\sum_{j=1}^{n} \boldsymbol{P}(|\hat{\theta}_{jn}-\theta_{jn}|[L(G_n, \hat{G}_n)\leqq\varepsilon])\leqq a_0\varepsilon$$

where $a_0=4m\{17+24m+(7+12m)(d-c)\}$.

PROOF. Fix n and $\boldsymbol{\theta}\in[c, d]^n$. We also fix j until (20) and X abbreviates X_j. Since $(3)-X'=\int_{X'}^{X}(\theta-X')q(\theta)dG_n(\theta)\Big/\int_{X'}^{X} q(\theta)dG_n(\theta)$, we abbreviate the quotient of the rhs to y/z and that with G_n replaced by $\hat{G}_n$ to Y/Z. Then,

$$(11)\qquad \hat{\theta}_{jn}-\theta_{jn}=\frac{Y}{Z}-\frac{y}{z}\ .$$

Let $*$ denote conditioning on X and $\{L(G_n, \hat{G}_n)\leqq\varepsilon\}$. Then, by Lemma 1.1 with $\gamma=1$ and $B=1$ and by the fact that $0\leqq Y/Z,\ y/z\leqq1$ we have

(12) $$P_*\left|\frac{Y}{Z}-\frac{y}{z}\right| \leqq \frac{2}{z} P_*(|Y-y|+2|Z-z|)\ .$$

By letting $I=(X', X]$, define by G_I the retraction of G into the closed interval $[G_n(X'), G_n(X)]$. Then, by Proposition (Nogami [7]),

$$L(G_I, \hat{G}_I) \leqq L(G_n, \hat{G}_n) \vee S \vee T$$

where $S=|(G_n-\hat{G}_n)(X')|$ and $T=|(G_n-\hat{G}_n)(X)|$. Thus,

(13) $$L(G_n, \hat{G}_n) \leqq \varepsilon \Rightarrow L(G_I, \hat{G}_I) \leqq \varepsilon \vee S \vee T \ (=\lambda)\ .$$

By applying Lemma 1.2 (Nogami [7]) with $h(\theta)$, the retraction of $(\theta-X')q(\theta)$ to I, and weakening the resulting bound,

(14) $$L(G_I, \hat{G}_I) \leqq \lambda \Rightarrow |Y-y| \leqq 2\alpha(\lambda+)+m(S+T)$$

where we use $+$ on the line to denote the right limit.

To bound $\alpha(\lambda+)$, pick $w_1, w_2 \in I$ such that $0<w_2-w_1<\lambda$. Now, by the definition of h,

(15) $$h(w_2)-h(w_1)=(w_2-w_1)\left\{q(w_2)+\frac{w_1-X'}{w_2-w_1}(q(w_2)-q(w_1))\right\}\ .$$

But, since by the definition of q, $q(w_2)-q(w_1)=q(w_2)q(w_1)\left(\int_{w_1}^{w_2} f(s)ds-\int_{w_1+1}^{w_2+1} f(s)ds\right)$ and since $q \leqq m$ and $0 \leqq f \leqq 1$,

(16) $$|q(w_2)-q(w_1)| \leqq m^2(w_2-w_1)\ .$$

Thus, from (15), $|h(w_2)-h(w_1)| \leqq (w_2-w_1)\{q(w_2)+(w_1-X')m^2\}$. Using $q \leqq m$, $w_1-X' \leqq 1$ and $w_2-w_1 \leqq \lambda$, and applying the definition of $\alpha(\lambda)$ gives us that

(17) $$\begin{aligned}\alpha(\lambda) &\leqq \vee\{h(w_2)-h(w_1)\colon \text{for } w_1, w_2 \in I \text{ such that } 0<w_2-w_1<\lambda\} \\ &\leqq \lambda(m+m^2)\end{aligned}$$

and thus the same bound applies for $\alpha(\lambda+)$.

Therefore, applying the bound of (17) to the first term of rhs (14) shows that

(18) $$\begin{aligned}L(G_I, \hat{G}_I) \leqq \lambda \Rightarrow |Y-y| &\leqq 2\lambda(m+m^2)+m(S+T) \\ &=2(\varepsilon+S+T)(m+m^2)+m(S+T)\ .\end{aligned}$$

Similarly, by Lemma 1.2 (Nogami [7]) with $1 \leqq h(=q) \leqq m$, when $L(G_I, \hat{G}_I) \leqq \lambda$, $|Z-z| \leqq 2\alpha(\lambda+)+m(S+T)$. Since by the definitions of $\alpha(\lambda)$ and q and by (16) $\alpha(\lambda) \leqq \vee\{|q(w_2)-q(w_1)|\colon \text{for } w_1, w_2 \in I \text{ such that } 0<w_2-w_1<\lambda\} \leqq m^2$, $\alpha(\lambda+)$ is also bounded by $m^2\lambda$. Hence, as in (18),

$$L(G_I, \hat{G}_I) \leqq \lambda \Rightarrow |Z-z| \leqq 2\lambda m^2 + m(S+T) . \tag{19}$$

Therefore, by (18), (19) and (13) and weakening the bound by replacing λ there by $\varepsilon + S + T$,

$$L(G_n, \hat{G}_n) \leqq \varepsilon \Rightarrow |Y-y| + 2|Z-z| \leqq 2(m+3m^2)\varepsilon + (5m+6m^2)(S+T) . \tag{20}$$

By this and in view of (12) and (11),

$$\begin{aligned} n^{-1} \sum_{j=1}^{n} \boldsymbol{P}(|\hat{\theta}_{jn} - \theta_{jn}|[L(G_n, \hat{G}_n) \leqq \varepsilon]) \\ \leqq 4(m+3m^2)\varepsilon\left(n^{-1} \sum_{j=1}^{n} P_{\theta_j} z^{-1}\right) + 2(5m+6m^2) \\ \times n^{-1} \sum_{j=1}^{n} \boldsymbol{P}\{(S+T)[L(G_n, \hat{G}_n) \leqq \varepsilon]/z\} . \end{aligned}$$

Applying Lemma 1.3 with $s=c-d-1$, $t=d+1-c$ and $\eta=\delta=0$ to the first term and using Lemma 1.4 twice to the second term results in the bound of the asserted lemma.

The following theorem is an immediate consequence of (6), (7) and Lemma 1.5.

THEOREM 1.1. *If $P_{\theta_j} \in \mathscr{P}(f)$ with $\Omega=[c, d]$, for $j=1, 2, \cdots, n$, then* $\varepsilon > 0$,

$$2^{-1}|D(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})| \leqq \boldsymbol{P}[L(G_n, \hat{G}_n) > \varepsilon] + a_0 \varepsilon \qquad \textit{uniformly in } \boldsymbol{\theta} ,$$

where a_0 is as defined in Lemma 1.5.

2. A particular procedure $\hat{\boldsymbol{\theta}}$ with a rate $O((n^{-1} \log n)^{1/4})$

We first construct a normalized (but not monotonized) estimate G_n^* of the empiric distribution function G_n. Then, we exhibit a distribution-valued estimate $\hat{G}_n$ of G_n. Main work in this section is, under the extra assumption on f (Lipschitz condition for $1/f$), to obtain the generalization (Lemma 2.5) of Lemma 3.1 of Fox [2]. Lemma 2.6 showing Lévy consistency of $\hat{G}_n$ to G_n, will be proved as in the proof of Theorem 3.1 of Fox [2] by using Lemma 2.5. Lemma 2.4 will be furnished to apply Hoeffding's bound ([5], Theorem 2) in the proof of Lemma 2.5. Finally, Theorem 2.1 shows that there exists a procedure $\hat{\boldsymbol{\theta}}$ with a rate $O((n^{-1} \log n)^{1/4})$.

In addition to the assumption on f in the introduction we now assume that $1/f$ satisfies the Lipschitz condition:

$$\vee \{(v-u)^{-1}|(f(v))^{-1} - (f(u))^{-1}| : u < v\} \leqq M \tag{21}$$

for a finite constant M. By this assumption,

$$|f(s)/f(t)-1|\leqq M|s-t| . \tag{22}$$

Let $\Omega=R$ through the end of the proof of Lemma 2.6. Let Q_n be the distribution function defined by

$$Q_n(\cdot)=\int_{-\infty}^{\cdot} qdG_n .$$

Then, letting $\bar{p}=\int p_\theta dG_n(\theta)$, we have by the definition of p_θ that $\bar{p}(y)=f(y)(Q_n(y)-Q_n(y'))$ and thus

$$Q_n(y)=\sum_{r=0}^{\infty}\frac{\bar{p}(y-r)}{f(y-r)} . \tag{23}$$

We remark that if the rth term of rhs(23) is nonzero then

$$r-1\leqq \text{range of } \{\theta_1,\cdots,\theta_n\} . \tag{24}$$

Since $q\geqq 1$ and q is the density of Q_n wrt G_n, it follows by Theorem 32.B of Halmos [4] that

$$G_n(y)=\int_{-\infty}^{y}(q(\theta))^{-1}dQ_n(\theta) . \tag{25}$$

For each y, we let $F_n^*(y)=n^{-1}\sum_{j=1}^{n}[X_j\leqq y]$ and for any $h>0$ $\Delta F_n^*(Y)=h^{-1}(F_n^*(y+h)-F_n^*(y))$. We allow h to depend on n and assume $h<1$ for convenience. Let $\bar{P}=\int P_\theta dG_n$. Then, $\bar{p}=d\bar{P}/d\xi$. We estimate $\bar{p}(y)$ by $\Delta F_n^*(y)$ and $Q_n(y)$ by

$$Q_n^*(y)=\sum_{r=0}^{\infty}(\Delta F_n^*(y-r)/f(y-r)) . \tag{26}$$

Note that Q_n^* has bounded variation because of (21). From (25), we obtain a raw estimate $\bar{W}$ of G_n from

$$\bar{W}(y)=\int_{-\infty}^{y}(q(t))^{-1}dQ_n^*(t) . \tag{27}$$

Since $F_n^*(y)\leqq G_n(y)\leqq F_n^*(y+1)$ for all $y\in R$, we furthermore estimate G_n at a point y by

$$G_n^*(y)=(F_n^*(y)\vee\bar{W}(y))\wedge F_n^*(y+1) .$$

We let $\delta=N^{-1}$, N being a positive integer depending on n, and consider the following grid on the real line: $\cdots<-2\delta<-\delta<0<\delta<2\delta<\cdots$. We finally estimate G_n at y by

(28) $$\hat{G}_n(y)=\sup\{G_n^*(j\delta):\ j\delta\leqq y,\ j=0,\pm 1,\cdots\}\ .$$

Let $\hat{\boldsymbol{\theta}}$ be the procedure whose component procedures are Bayes versus $\hat{G}_n$ given by (28). To get a rate of convergence of the modified regret for $\hat{\boldsymbol{\theta}}$ we use the bound of Theorem 1.1; we shall get an upper bound for $\boldsymbol{P}[L(G_n, \hat{G}_n)>2\varepsilon]$ (forthcoming Lemma 2.6). This bound is essentially given by bounding $1-\mathrm{P}(\{G_n(y-\varepsilon)-\varepsilon\leqq G_n^*(y)\leqq G_n(y+\varepsilon)+\varepsilon\})$ (forthcoming Lemma 2.5). Therefore, the main part in this section is Lemma 2.5. But, since we will apply Theorem 2 of Hoeffding [5] for its proof, we want to get the bounds for $\boldsymbol{P}\bar{W}(y)$; an upper bound with the terms $G_n(y+h)$ and $o(h)$ and a lower bound with the terms $G_n(y)$ and $o(h)$ (Lemma 2.4). To do so we shall furnish Lemmas 2.1, 2.2 and 2.3.

Since the summation of r in (26) involves at most a finite number of non-zero terms, we shall freely interchange integral and summation on r without further comment.

By the definition (27) of $\bar{W}$, $\bar{W}=n^{-1}\sum_{j=1}^{n} W_j$ where for each j

(29) $$W_j=\sum_{r=0}^{\infty}\int_{-\infty}^{y}(q(t))^{-1}d_t\{[t-r<X_j\leqq t-r+h](hf(t-r))^{-1}\}$$

where the subscript t in d_t denotes the variable of integration. To find bounds of $\boldsymbol{P}W(y)$ we shall find an upper and lower bound of $\boldsymbol{P}W_j$, for each j. Fix j and use the corresponding notations without subscript j until the end of the proof of Lemma 2.4. We shall start with getting an alternative form of $\boldsymbol{P}W$. Hereafter, we abbreviate $f(b)-f(a)$ to $f]_a^b$ until the end of the proof of Lemma 2.3.

LEMMA 2.1.

(30) $$\boldsymbol{P}(W)/q(\theta)=(q(y))^{-1}\sum_{r=0}^{\infty}S(y-r)-I(S)$$

where

(31) $$S(t)=h^{-1}\int_{t}^{t+h}[\theta\leqq s<\theta+1](f(s)/f(t))ds$$

and

$$I(S)=\int_{-\infty}^{y}\sum_{r=0}^{\infty}S(t-r)f]_t^{t+1}dt\ .$$

PROOF. Because a function satisfying the Lipschitz condition is absolutely continuous (cf. Royden [10], p. 108) and $1/q$ is clearly absolutely continuous, $1/f(\cdot\,-r)$ and $1/q$ are both of bounded variation in the definition (29) of W. Applying integration by parts (Saks [11], Theorem III.14.1) and using $d(q(t))^{-1}=f]_t^{t+1}dt$ gives us that

$$(32)\quad \int_{-\infty}^{y}(q(t))^{-1}d_t([t-r<X\leqq t-r+h]/f(t-r))$$
$$=\frac{[y-r<X\leqq y-r+h]}{f(y-r)q(y)}-\int_{-\infty}^{y}\frac{[t-r<X\leqq t-r+h]}{f(t-r)}f]_t^{t+1}dt\ .$$

Now, with $\mathrm{E}\,X$ denoting the expectation of a random variable X, Proposition III.2.1 of Neveu [6] gives us a version of the relation $\mathrm{E}\{\mathrm{E}(h(t)|X)\}=\mathrm{E}\,h(t)$ for an integrable function h and probability measures. But, because of its proof it holds for finite measures. Hence,

$$P_\theta\left\{\int_{-\infty}^{y}\frac{[t-r<X\leqq t-r+h]}{f(t-r)}f]_t^{t+1}dt\right\}=hq(\theta)\int_{-\infty}^{y}S(t-r)f]_t^{t+1}dt\ .$$

Thus, taking expectation wrt X and then summation on r over (32) and multiplying $(hq(\theta))^{-1}$ on both sides gives us the asserted lemma.

To get bounds for $\boldsymbol{PW}$ we shall first find bounds for the first term of rhs (30) (Lemma 2.2) and then bounds for the second term $I(S)$ of rhs (30) (Lemma 2.3). Until the end of the proof of Lemma 2.3 we use the notation

$$\Delta(t)=h^{-1}\int_{t}^{t+h}[\theta\leqq s<\theta+1]ds\ .$$

LEMMA 2.2.

$$(33)\quad (q(y))^{-1}[\theta\leqq y]-2^{-1}Mh\leqq \textit{first term of rhs}\,(30)$$
$$\leqq(q(y))^{-1}[\theta\leqq y+h]+2^{-1}Mh\ .$$

PROOF. Applying (22) to the definition (31) of $S(t)$ and changing a variable leads to the inequality

$$(34)\quad |S(y-r)-\Delta(y-r)|\leqq Mh^{-1}\int_0^h[\theta-(y-r)\leqq u<\theta+1-(y-r)]udu\ .$$

Moreover, because

$$(35)\quad \sum_{r=0}^{\infty}\Delta(y-r)=h^{-1}\int_{y}^{y+h}[\theta\leqq t]dt=h^{-1}[\theta-h\leqq y<\theta](y+h-\theta)+[\theta\leqq y]\ ,$$

we obtain that $[\theta\leqq y]\leqq\sum_{r=1}^{\infty}\Delta(y-r)\leqq[\theta\leqq y+h]$. Hence, this, $\sum_{r=0}^{\infty}\mathrm{rhs}\,(34)\leqq 2^{-1}Mh$ and $(q(y))^{-1}\leqq 1$ leads to the bounds in the asserted lemma.

LEMMA 2.3.

$$(36)\quad \left|I(S)-[\theta\leqq y]\left(\int_\theta^y f]_t^{t+1}dt\right)\right|\leqq(M+M\wedge 1)2^{-1}h\ .$$

PROOF. Let $[z]$ denote the greater integer $\leqq z$ if $z>0$ and -1 if

$z<0$. Since $\int_{-\infty}^{y} S(t-y)f]_t^{t+1}dt=\int S(t)[t\leqq y-r]f]_{t+r}^{t+r+1}dt$ and

$$(37)\qquad \sum_{r=0}^{\infty}[t\leqq y-r](f(t+r+1)-f(t+r))$$
$$=[t\leqq y](f(t+[y-t]+1)-f(t))=f(t+[y-t]+1)-f(t)$$

(the latter equality because $[y-t]=-1$ if $t>y$), if follows that

$$(38)\qquad I(S)=\int S(t)f]_t^{t+[y-t]+1}dt\ .$$

From the derivation, (38) holds for Δ in place of S. Thus, from (34) with any $y-r$ and then by $0\leqq f\leqq 1$,

$$(39)\qquad |I(S)-I(\Delta)|\leqq 2^{-1}Mh\ .$$

But by (35), $I(\Delta)=\int_{-\infty}^{y}(\text{lhs}\,(35)\text{ with } y=t)f]_t^{t+1}dt$ which in turn equals

$$(40)\qquad \left\{[\theta-h\leqq y<\theta]\int_{\theta-h}^{y}+[\theta\leqq y]\int_{\theta-h}^{\theta}\right\}h^{-1}(t+h-\theta)f]_t^{t+1}dt$$
$$+[\theta\leqq y]\left(\int_{\theta}^{y}f]_t^{t+1}dt\right)\ .$$

Since $|f(t+1)-f(t)|\leqq|(f(t))^{-1}-(f(t+1))^{-1}|\wedge 1\leqq M\wedge 1$ and $\int_{\theta-h}^{y\wedge\theta}(t+h-\theta)dt\leqq 2^{-1}h^2$, |first term of (40)|$\leqq(M\wedge 1)2^{-1}h$. Thus,

$$\left|I(\Delta)-[\theta\leqq y]\left(\int_{\theta}^{y}f]_t^{t+1}dt\right)\right|\leqq(M\wedge 1)2^{-1}h\ ,$$

and hence by (39) we get the asserted bound of the lemma.

LEMMA 2.4. *For every* $y\in[\theta_{(1)}-1,\theta_{(n)}+1]$,

$$(41)\qquad G_n(y)-b_1h\leqq \boldsymbol{P}\bar{W}(y)\leqq G_n(y+h)+b_1h$$

where $b_1=2^{-1}m(2M+5(M\wedge 1))$, $\theta_{(1)}=\min_{1\leqq i\leqq n}\theta_i$ and $\theta_{(n)}=\max_{1\leqq i\leqq n}\theta_i$.

PROOF. We shall first find the bounds of $\boldsymbol{P}W$. From (30), (33) and (36), we can see that $(\boldsymbol{P}W)/q(\theta)$ is bounded above and below by rhs(33)−(lower bound of $I(S)$ in (36)) and lhs(33)−(upper bound of $I(S)$ in (36)), respectively.

Since for $h>0$

$$(42)\qquad [\theta\leqq y]\int_{\theta}^{y}f]_t^{t+1}dt=[\theta\leqq y+h]/q(y)-[\theta\leqq y+h]/q(\theta)$$
$$-[y<\theta\leqq y+h]\int_{y}^{\theta}f]_{t+1}^{t}dt\ ,$$

using (42) and weakening the upper bound for $(\boldsymbol{P}W)/q(\theta)$ by applying

$$[y<\theta\leqq y+h]\left|\int_y^\theta f]_{t+1}^t dt\right|\leqq(M\wedge 1)h \tag{43}$$

and $1\leqq q\leqq m$ results in the upper bound

$$\boldsymbol{P}W\leqq[\theta\leqq y+h]+b_1h-(M\wedge 1)h\leqq[\theta\leqq y+h]+b_1h\ .$$

On the other hand, since applying the equality $[y<\theta\leqq y+h]\{(q(\theta))^{-1}-(q(y))^{-1}\}=-\{$the third term of rhs (42)$\}$ results in the equality lhs (42) $=[\theta\leqq y]\{(q(y))^{-1}-(q(\theta))^{-1}\}-2\{-($the third term of rhs (42)$)\}$, weakening the lower bound for $(\boldsymbol{P}W)/q(\theta)$ by using (43) and $q\leqq m$ we obtain

$$[\theta\leqq y]-b_1h\leqq\boldsymbol{P}W.$$

Averaging the above upper and lower bounds for $\boldsymbol{P}W$ gives the asserted bounds for the lemma.

Following Lemma 2.5 is a direct generalization of Lemma 3.1 of Fox [2] in the sense that if $f\equiv 1$, then $m=1$ and $M=0$, and hence we get his bound $2\exp(-2nh^2\varepsilon^2)$.

LEMMA 2.5. *If* $0<h\leqq\varepsilon\leqq 1$, *then for each* y

$$\begin{aligned}1-\mathrm{P}(\{G_n(y-\varepsilon)-\varepsilon\leqq G_n^*(y)\leqq G_n(y+\varepsilon)+\varepsilon\})\\ \leqq 2\exp\left\{-\frac{2nh^2((\varepsilon-b_1h)^+)^2}{(1+3b_0M)^2}\right\}\end{aligned} \tag{44}$$

where $b_0=d-c+3$ *and* b_1 *is as defined in Lemma* 2.4.

PROOF. For $y>\theta_{(n)}+1$, $F_n^*(y)=G_n^*(y)=G_n(y+\varepsilon)=1$ and for $y<\theta_{(1)}-1$, $F_n^*(y+1)=G_n^*(y)=G_n(y-\varepsilon)=0$; in both case lhs $(44)=0$ and (44) holds trivially.

For $y\in[\theta_{(1)}-1,\theta_{(n)}+1]$, it is sufficient to prove the lemma for the raw estimate $\bar{W}$, for if $G_n(y-\varepsilon)-\varepsilon\leqq\bar{W}(y)\leqq G_n(y+\varepsilon)+\varepsilon$, it follows that $G_n(y-\varepsilon)-\varepsilon\leqq\bar{W}(y)\wedge F_n^*(y+1)\leqq G_n^*(y)\leqq\bar{W}(y)\vee F_n^*(y)\leqq G_n(y+\varepsilon)+\varepsilon$.

Pick $y\in[\theta_{(1)}-1,\theta_{(n)}+1]$. As in the proof of Lemma 3.1 of Fox [2] we shall apply Theorem 2 of Hoeffding [5]. To do so we shall use the bounds of $\boldsymbol{P}(\bar{W}(y))$ in Lemma 2.4 and furthermore need to get an upper and a lower bound of W_j, for each j. By (32) and (37) applied to the definition (29) of W_j

$$\begin{aligned}hW_j=(q(y))^{-1}\sum_{r=0}^{\infty}[y-r<X_j\leqq y-r+h]/f(y-r)\\ -\int[t<X_j\leqq t+h]\{(f(t+[y-t]+1)/f(t))-1\}dt\ .\end{aligned} \tag{45}$$

In the summation of the first term of rhs(45), there are at most two positive terms and both terms cannot be positive at the same time. Applying (22) and then the fact that $r\leqq d-c+2$ $(=b_0-1)$ gives that

$$0\leqq(\text{first term of rhs}(45))\leqq 1+b_0M\,.$$

In addition, by a use of (22) and the fact that $[y-X_j+h]\leqq b_0-1$ (because $y\leqq\theta_{(n)}+1$, $\theta_{(1)}\leqq X_j$ and $h<1$),

$$|\text{second term of rhs}(45)|\leqq b_0Mh \qquad (<b_0M)\,.$$

Therefore,

$$-b_0M\leqq hW_j\leqq 1+2b_0M\,, \qquad \text{for each } j\,.$$

We now apply Theorem 2 of Hoeffding [5]. Since $h\leqq\varepsilon$, using the second inequality of (41) in Lemma 2.4 and applying Theorem 2 of Hoeffding [5] gives

$$\begin{aligned}(46)\qquad \boldsymbol{P}[\bar{W}(y)>G_n(y+\varepsilon)+\varepsilon]&\leqq\boldsymbol{P}[\bar{W}(y)-\boldsymbol{P}\bar{W}(y)>\varepsilon-b_1h]\\ &\leqq\exp\left\{-\frac{2nh^2((\varepsilon-b_1h)^+)^2}{(1+3b_0M)^2}\right\}\,.\end{aligned}$$

Furthermore, by the first inequality of (41), $\{\bar{W}(y)<G_n(y-\varepsilon)-\varepsilon\}\subset\{\boldsymbol{P}\bar{W}(y)-\bar{W}(y)>\varepsilon-b_1h\}$. Hence by the symmetry of the tail bounds, $\boldsymbol{P}[\bar{W}(y)<G_n(y-\varepsilon)-\varepsilon]\leqq\text{rhs}(46)$, which together with (46) gives us the asserted bound of Lemma 2.5.

With $\hat{G}_n$ as defined in (28) we are now furnished to get an upper bound for the first term of the right hand side of Theorem 1.1, showing Lévy consistency of the estimate $\hat{G}_n$ for G_n.

LEMMA 2.6. *For any $\varepsilon>0$, if $h\leqq\varepsilon$ and $\delta\leqq\varepsilon$, then*

$$(47)\qquad \boldsymbol{P}[L(G_n,\hat{G}_n)>2\varepsilon]\leqq(\delta^{-1}+1)[\varepsilon^{-1}+1](\text{rhs}(44))\,.$$

PROOF. We rely on the proof of Theorem 3.1 of Fox [2]. For $0<\varepsilon\leqq1$, let n be large enough so that $h\leqq\varepsilon$ and $\delta\leqq\varepsilon$. Let J be the largest integer such that $F_n^*(j\delta+1)\leqq\varepsilon$. We also let $\mathcal{T}=\{j\colon F_n^*((j+1)\delta+1)-F_n^*(j\delta)>\varepsilon,\ j\geqq J,\ j=0,\pm1,\cdots\}$ and $A_n=\bigcup_{j\in\mathcal{T}}[j\delta,(j+1)\delta)$. Since only retraction and monotonicity properties of his respective estimate G_n^* and G_n were used before Lemma 3.1 of Fox was applied, the following inequalities are still true for our estimates G_n^* and $\hat{G}_n$.

$$\begin{aligned}(48)\qquad &\boldsymbol{P}[L(G_n,\hat{G}_n)>2\varepsilon]\\ &=\boldsymbol{P}\Big(\bigcup_{y\in A_n}(\{\hat{G}_n(y)>G_n(y+2\varepsilon)+2\varepsilon\}\cup\{\hat{G}_n(y)<G_n(y-2\varepsilon)-2\varepsilon\})\Big)\end{aligned}$$

$$\leqq \boldsymbol{P} \bigcup_{j\delta \in A_n} (\{G_n^*(j\delta) > G_n(j\delta+\varepsilon)+\varepsilon\} \cup \{G_n^*(j\delta) < G_n(j\delta-\varepsilon)-\varepsilon\})$$
$$\leqq \sum_{j\delta \in A_n} \boldsymbol{P}(\{G_n^*(j\delta) > G_n(j\delta+\varepsilon)+\varepsilon\} \cup \{G_n^*(j\delta) < G_n(j\delta-\varepsilon)-\varepsilon\}) .$$

Since there are at most $(\delta^{-1}+1)[\varepsilon^{-1}+1]$ grid points (see Fox [2], p. 1850) in A_n, by Lemma 2.5 the extreme rhs (48) is no larger than rhs (47).

Let $\hat{\boldsymbol{\theta}}$ be the procedure whose component procedures are Bayes versus $\hat{G}_n$ defined by (28). To get a rate of convergence of the modified regret for $\hat{\boldsymbol{\theta}}$ we use the bound of Theorem 1.1. Since this bound is valid only for $\Omega=[c, d]$ where $-\infty<c\leqq d<+\infty$, we consider only $\mathscr{P}(f)$ with $\Omega=[c, d]$.

THEOREM 2.1. *If $P_{\theta_j} \in \mathscr{P}(f)$ with $\Omega=[c, d]$, $j=1, 2, \cdots, n$ where f^{-1} satisfies the Lipschitz condition* (21), *then there exist constants b_2 and b_3 so that, for $\hat{\boldsymbol{\theta}}$ with $b_2 h=b_3\delta=(n^{-1}\log n)^{1/4}$,*

$$|D(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})|=O((n^{-1}\log n)^{1/4}) , \qquad \textit{uniformly in } \boldsymbol{\theta} \in [c, d]^n . \tag{49}$$

PROOF. We use Theorem 1.1 and apply Lemma 2.6. Then, choosing $\varepsilon=\delta=(2b_1+1)h<1$ (for sufficiently large n) and weakening the bound gives

$$|D(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})|\leqq b_4 h+b_5 h^{-2} \exp\{-(nh^4/b_6)\} \tag{50}$$

where b_4 and b_5 are some constants, and $b_6=2\{1+3(d-c+3)M\}^2$.

Choose b_2 and b_3 so that $b_2\leqq 4^{1/4}(3b_6)^{-1/4}$ and $b_3=b_2(2b_1+1)^{-1}$. Then, for $b_2 h(=b_3\delta)=(n^{-1}\log n)^{1/4}$, (50) leads to the asserted rate in Theorem 2.1.

3. A counterexample to $D(\boldsymbol{\theta}, \boldsymbol{t})\to 0$ on R^∞

In Section 2 we demonstrated a procedure $\hat{\boldsymbol{\theta}}$ such that $|D(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})|= O((n^{-1}\log n)^{1/4})$ uniformly in $\boldsymbol{\theta}$ in case of a bounded parameter set $\Omega= [c, d]$. Here we prove that the boundedness assumption on Ω is necessary for the modified regret converging to zero.

THEOREM 3.1. *Let $X_1, X_2, \cdots$ be independent random variables where for each j, $X_j\sim U[\theta_j, \theta_j+1)$, $\theta_j \in \Omega=R$. Let $\boldsymbol{t}(\boldsymbol{X})=(t_1(\boldsymbol{X}), \cdots, t_n(\boldsymbol{X}))$ be an estimator of $\boldsymbol{\theta}=(\theta_1, \cdots, \theta_n)$, $n=1, 2, \cdots$. Then, there exists a sequence $(\theta_1, \theta_2, \cdots) \in R^\infty$ such that $\overline{\lim}_{n\to\infty} D(\boldsymbol{\theta}, \boldsymbol{t})>0$.*

PROOF. $\boldsymbol{P}_x$ denotes the conditional distribution of $(X_1, \cdots, X_{j-1}, X_{j+1}, \cdots, X_n)$ given $x=X_j$. Since for each j, $\boldsymbol{P}(t_j(\boldsymbol{X})-\theta_j)^2\geqq P_{\theta_j}(\boldsymbol{P}_x(t_j(\boldsymbol{X}))$

$-\theta_j)^2$, it follows that

$$D(\boldsymbol{\theta}, \boldsymbol{t}) \geqq n^{-1} \sum_{j=1}^{n} P_{\theta_j}(\boldsymbol{P}_x(t_j(\boldsymbol{X}))-\theta_j)^2 - R(G_n) . \tag{51}$$

Now, let $\boldsymbol{\mu}$ be a joint prior measure on $(\theta_1, \theta_2, \cdots)$. Let $\boldsymbol{\mu}_{\theta_j}$ be the conditional measure given θ_j and let μ_j be the marginal measure of θ_j. Then, setting $s_j = \boldsymbol{\mu}_{\theta_j}\boldsymbol{P}_x(t_j(\boldsymbol{X}))$, $j=1, 2, \cdots, n$, we have that

$$\boldsymbol{\mu}\left\{n^{-1} \sum_{j=1}^{n} P_{\theta_j}(\boldsymbol{P}_x(t_j(\boldsymbol{X}))-\theta_j)^2\right\} \geqq n^{-1} \sum_{j=1}^{n} \mu_j P_{\theta_j}(s_j-\theta_j)^2 . \tag{52}$$

Now consider $\boldsymbol{\mu} = \mu_1 \times \mu_2 \times \cdots$ where μ_j puts mass 1/2 on each of the values $2j \pm r$, $j \geqq 1$, where r is some fixed number such that $0 < r < 1/2$. Then

$$\begin{aligned} \mu_j P_{\theta_j}(s_j-\theta_j)^2 &= 2^{-1}P_{2j-r}(s_j-(2j-r))^2 + 2^{-1}P_{2j+r}(s_j-(2j+r))^2 \\ &\geqq \int_{2j+r}^{2j+1-r} \{2^{-1}(s_j-(2j-r))^2 + 2^{-1}(s_j-(2j+r))^2\}\,dx \\ &\geqq r^2(1-2r) , \end{aligned} \tag{53}$$

where the last inequality follows since the integrand on the lhs is not less than r^2.

Since $R(G_n) = n^{-1} \sum_{j=1}^{n} P_{\theta_j}(\theta_{jn}-\theta_j)^2$ where θ_{jn} is defined by the posterior mean (3) with $q \equiv 1$, and since the θ_j's are apart from each other more than 1, $\theta_{jn} = \theta_j$ for all j and hence $R(G_n)=0$. Thus, $\boldsymbol{\mu}(R(G_n))=0$. Therefore, in view of (51), (52) and (53),

$$\boldsymbol{\mu}\{D(\boldsymbol{\theta}, \boldsymbol{t})\} \geqq r^2(1-2r) \tag{54}$$

for all n. The retraction $\boldsymbol{t}^*$ of $\boldsymbol{t}$ formed by taking $t_j^* = (X_j' \wedge t_j) \vee X_j$ has modified regret bounded by 1 and satisfies (54). Therefore, using Fatou's lemma gives

$$\boldsymbol{\mu}\{\overline{\lim_{n\to\infty}}\, D(\boldsymbol{\theta}, \boldsymbol{t}^*)\} \geqq \overline{\lim_{n\to\infty}}\, \{\boldsymbol{\mu} D(\boldsymbol{\theta}, \boldsymbol{t}^*)\} \geqq r^2(1-2r) > 0 . \tag{55}$$

By $\overline{\lim}_{n\to\infty} D(\boldsymbol{\theta}, \boldsymbol{t}) \geqq \overline{\lim}_{n\to\infty} D(\boldsymbol{\theta}, \boldsymbol{t}^*)$ and (55), there exists a $(\theta_1, \theta_2, \cdots) \in R^\infty$ such that $\overline{\lim}_{n\to\infty} D(\boldsymbol{\theta}, \boldsymbol{t}) > 0$.

Acknowledgment

The author heartily express thanks to Professor James Hannan for suggesting Lemma 1.2 and a counterexample in Section 3 and for his excellent guidance during this research work. The author is also thankful to the referee for valuable comments to refine the manuscript.

THE UNIVERSITY OF TSUKUBA

REFERENCES

[1] Feller, W. (1971). *An Introduction to Probability Theory and its Applications Volume* II (2nd ed.), Wiley, New York.

[2] Fox, R. (1970). Estimating the empiric distribution function of certain parameter sequences, *Ann. Math. Statist.*, **41**, 1845-1852.

[3] Fox, R. (1978). Solutions to empirical Bayes squared error loss estimation problems, *Ann. Statist.*, **6**, 846-853.

[4] Halmos, P. R. (1950). *Measure Theory*, Litton Educational Publishing.

[5] Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables, *J. Amer. Statist. Ass.*, **58**, 13-30.

[6] Neveu, J. (1965). *Mathematical Foundations of the Calculus of Probability*, Holden-Day.

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

[8] Nogami, Y. (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.

[9] Oaten, A. (1969). Approximation to Bayes risk in compound decision problems, RM-233, Department of Statistics and Probability, Michigan State University.

[10] Royden, H. L. (1968). *Real Analysis*, The Macmillan Company, New York.

[11] Saks, S. (1937). *Theory of the Integral* (2nd ed.), Monografie Mathematyczne.

[12] Singh, Radhey S. (1977). Sequence-compound estimation with rate in non-continuous Lebesgue-exponential families, *Sankhyā*, **39**, A, 63-81.