Ann. Inst. Statist. Math.
33 (1981), Part A, 67-80

# THE SET-COMPOUND ONE-STAGE ESTIMATION IN THE NONREGULAR* FAMILY OF DISTRIBUTIONS OVER THE INTERVAL $[\theta, \theta+1)$

YOSHIKO NOGAMI

(Received Apr. 6, 1978; revised Nov. 20, 1980)

## Introduction

This paper is a continuation of author's Ph.D. thesis [6] and Nogami [7].

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

Let $\xi$ be Lebesgue measure and $f$ a measurable function with $0 \leqq f \leqq 1$. Define

$$q(\theta)=1\Big/\int_{\theta}^{\theta+1} f d\xi\,. \tag{0.1}$$

Let $\mathcal{P}^*(f)$ be a family of probability measures determined by

$$\mathcal{P}^*(f)=\{P_\theta \text{ with } p_\theta=q(\theta)[\theta, \theta+1)f,\ \forall\, \theta \in \Omega\} \tag{0.2}$$

where $p_\theta=dP_\theta/d\xi$ and $\Omega$ is a real interval $[c, d]$ with $-\infty<c<d<+\infty$, and we denote the indicator function of a set $A$ by $A$ itself. The component problem is the squared-error loss estimation (SELE) of $\theta$ based on $X \sim P_\theta \in \mathcal{P}^*(f)$.

Let $X_1, \cdots, X_n$ be $n$ independent random variables with each $X_j \sim P_{\theta_j} \in \mathcal{P}^*(f)$. The modified regret of the set-compound decision procedure $\boldsymbol{t}=(t_1, \cdots, t_n)$ is of form

$$D(\boldsymbol{\theta}, \boldsymbol{t})=\mathrm{E}\left(n^{-1}\sum_{j=1}^{n}(\theta_j-t_j(\boldsymbol{X}))^2\right)-R(G_n) \tag{0.3}$$

where $R(G_n)$ is the Bayes risk against the empiric distribution $G_n$ of $\theta_1, \cdots, \theta_n$ in the component problem.

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

* The word "nonregular" was quoted from Ferguson ([3], p. 130).

$$\theta_{jn} = \int_{X'_j+}^{X_j} \theta q(\theta) dG_n(\theta) \Big/ \int_{X'_j+}^{X_j} q dG_n \tag{0.4}$$

where $y'$ is an abbreviation of $y-1$ and the affix $+$ is intended to describe the integration as over $(X'_j, X_j]$. Henceforth we delete $+$ in lower limits of $\int$s. Then, we can write

$$R(G_n) = \mathrm{E}\left(n^{-1} \sum_{j=1}^{n} (\theta_j - \theta_{jn})^2\right).$$

The work here is a generalization and an extension of Fox's work [4], respective to a family $\mathcal{P}^*(f)$ and to the set-compound SELE problem. When $P_\theta$ is the uniform distribution on $[\theta, \theta+1)$ for $\theta \in (-\infty, \infty)$ and the $\theta$ are i.i.d. with a prior $G$, Fox [4] showed the convergence to $R(G)$ of the respective expected risks for a one-stage procedure with components direct estimates of the posterior means wrt $G$.

Nogami ([6], Chapters II and III) introduced a one-stage set-compound estimate $\boldsymbol{\theta}_T$ for $\boldsymbol{\theta} \in [c, d]^n$ (we say this fact so that $\boldsymbol{\theta}_T$ has a rate 1/4) under $\mathcal{P}^*(f)$ with Lipshitz condition for $1/f$. In this paper we demonstrate (in Section 1) another estimate $\boldsymbol{\phi}^*$ with a rate 1/3 without Lipshitz condition for $1/f$ and can expect (from Section 2) that both $\boldsymbol{\theta}_T$ and $\boldsymbol{\phi}^*$ have the same best exact order $n^{-2/3}$ of convergence of the modified regret. In Section 1 we get an upper bound for $D(\boldsymbol{\theta}, \boldsymbol{\phi}^*)$. In Nogami [8], there is a misprint in the bound of Theorem in Section 2. The bound there should be $(8N+24)m^k\{N^k \cdot (4k-2+k^{1/2})((n-k+1)h^k)^{-1/2}+2^{-k+1}h^k\}$. Although this bound with $k=1$ gives an upper bound for $D(\boldsymbol{\theta}, \boldsymbol{\phi}^*)$, the result of Section 1 in this paper is stronger than that. Section 2 gives us lower bounds for $D(\mathbf{0}, \boldsymbol{\phi}^*)$ at $f \equiv 1$.

*Notations.* We often let $P(h)$ or $P(h(\omega))$ denote $\int h(\omega) dP(\omega)$. $G$ abbreviates the empiric distribution $G_n$ of $\theta_1, \cdots, \theta_n$. For any function $h$, $h]_a^b$ means $h(b)-h(a)$. When we refer to (a. b) in Section a we simply write (b). $\vee$ and $\wedge$ denote the supremum and the infimum, respectively. $\doteq$ denotes the defining property. $\boldsymbol{P}_x$ means the conditional expectation of $X_1, \cdots, X_{j-1}, X_{j+1}, \cdots, X_n$ given $X_j \doteq x$. A distribution function also represents the corresponding measure. Define $\bar{z} = = n^{-1} \sum_{i=1}^{n} z_i$.

## 1. An upper bound for the modified regret $D(\boldsymbol{\theta}, \boldsymbol{\phi}^*)$

In this section we shall get a one-stage procedure $\boldsymbol{\phi}^*$ for estimating $\boldsymbol{\theta}$ and show that it has a rate 1/3 under the assumption $f(\cdot) \geqq m^{-1}$.

Assume

(1.1) $$f(\cdot)\geqq m^{-1}\ (>0)\,, \qquad \text{for a constant } m\ (>1)\,.$$

Before mentioning about the structure of the procedure $\boldsymbol{\phi}^*$ we introduce the following:

LEMMA 1.1 (Nogami [6]). *Let $\tau$ be a signed measure, $g$ be a measurable function and $I=(y', y]$ be an interval with $\int Igd\tau\neq 0$. Let $\tau_y$ be the signed measure with density $Ig/\int Igd\tau$ wrt $\tau$. Then,*

$$\int sd\tau_y(s)=y-\int_0^1 \tau_y(y', y'+t]dt\,.$$

PROOF. By Fubini's theorem applied to the lhs of the second equality below,

$$y-\int sd\tau_y(s)=\int\int_{s-y'}^1 dtd\tau_y(s)=\int_0^1 \tau_y(y', y'+t]dt\,.$$

For fixed $j$, $1\leqq j\leqq n$, we abbreviate $X_j$ to $x$. Fix $j$ until (5). Let $Q$ be the measure with the density $q$ wrt $G$. Define

(1.2) $$u_i(y)=p_i(y)/f(y)\,, \qquad i=1, \cdots, n\,.$$

Then, by the definition of $p_\theta$

(1.3) $$\bar{u}(y)=Q]_{y'}^{y}\,.$$

Thus,

(1.4) $$Q(y)=\sum_{r=0}^{\infty}\bar{u}(y-r)\,.$$

By above Lemma 1.1 applied to (0.4)

(1.5) $$\begin{aligned}\theta_{jn}&=x-\int_0^1 Q]_{x'}^{x'+t}dt/Q]_{x'}^{x}\\&=x-\int_0^1\sum_{r=1}^{\infty}\bar{u}]_{x-r}^{x-r+t}dt/\bar{u}(x)\,.\end{aligned}$$

In view of (2) we estimate $\bar{u}(y)$ by $\bar{\hat{u}}(y)=n^{-1}\sum_{i=1}^{n}\hat{u}_i$ where for any $h>0$

(1.6) $$\hat{u}_i(y)=h^{-1}[y\leqq X_i<y+h]/f(X_i)\,.$$

We allow $h$ to depend on $n$ and assume $h<1$ for convenience. Thus, this and (5) (observe $x'<\theta_{jn}\leqq x$) suggest that to achieve a small modified regret we might estimate $\theta_{jn}$ by

(1.7) $$\phi_{jn}^*=x-0\vee\left(\int_0^1\sum_{r=1}^{\infty}\bar{\hat{u}}]_{x-r}^{x-r+t}dt/\bar{\hat{u}}(x)\right)\wedge 1$$

and thus $\boldsymbol{\phi}^*=(\phi_{1n}^*, \cdots, \phi_{nn}^*)$ is an estimate of $\boldsymbol{\theta}_G(\boldsymbol{X})$ and thus of $\boldsymbol{\theta}$. Note that if the $r$-th term of the numerator of the quotient of rhs (7) is nonzero, then

$$r \leqq d+3-c \doteq N+2 . \tag{1.8}$$

Since $X_j' < \phi_{jn}^*$, $\theta_{jn} \leqq X_j$, $j=1, 2, \cdots, n$,

$$n2^{-1}|D(\boldsymbol{\theta}, \boldsymbol{\phi}^*)| \leqq \sum_{j=1}^{n} P_j \boldsymbol{P}_x |\phi_{jn}^* - \theta_{jn}| . \tag{1.9}$$

We shall invoke the following corollary, a special case of Lemma A.2 of Singh [9], to get a bound of $\boldsymbol{P}_x|\phi_{jn}^*-\theta_{jn}|$.

COROLLARY 1.1 (Singh [9]). *For real random variables $Y$ and $Z$, and real numbers $y$ and $z$,*

$$\mathrm{E}\left(\left|\frac{Y}{Z}-\frac{y}{z}\right| \wedge 1\right) \leqq 2|z|^{-1}\left\{\mathrm{E}\,|Y-y| + \left(\left|\frac{y}{z}\right|+1\right)\mathrm{E}\,|Z-z|\right\} . \tag{1.10}$$

Applying Corollary 1.1 and weakening the resulted bound shows that for fixed $j$,

$$\begin{aligned} \boldsymbol{P}_x|\phi_{jn}^*-\theta_{jn}| \leqq (\bar{u}(x))^{-1}\Bigg[\sum_{r=1}^{N+2}\Bigg\{\int_0^1 \boldsymbol{P}_x|\bar{u}(x-r+t)-\bar{\hat{u}}(x-r+t)|\,dt \\ +\boldsymbol{P}_x|\bar{u}(x-r)-\bar{\hat{u}}(x-r)|\Bigg\} + 2\boldsymbol{P}_x\,|\bar{u}(x)-\bar{\hat{u}}(x)|\Bigg] . \end{aligned} \tag{1.11}$$

But with $\bar{u}_j \doteq (n-1)^{-1} \sum_{(j\neq)i=1}^{n} u_i$ and $\bar{\hat{u}}_j \doteq (n-1)^{-1} \sum_{(j\neq)i=1}^{n} \hat{u}_i$,

$$\begin{aligned} n\boldsymbol{P}_x|\bar{u}(x-r+t)-\bar{\hat{u}}(x-r+t)|-(n-1)\boldsymbol{P}_x|\bar{u}_j(x-r+t)-\bar{\hat{u}}_j(x-r+t)| \\ \leqq |u_j(x-r+t)-(hf(x))^{-1}| \leqq (2m)/h \end{aligned} \tag{1.12}$$

where the last inequality follows by

$$u_j(\cdot) \leqq m , \quad \forall j \quad \text{and} \quad 1/f(\cdot) \leqq m . \tag{1.13}$$

Lemma 2.2 below will be used to get a bound of $\sum_{j=1}^{n} P_j(\text{rhs}\,(11))$ and is proved in the proof of Lemma 2.1 of Nogami [7] with $\beta$ there replaced by $N$.

LEMMA 2.2.

$$\sum_{j=1}^{n} P_j(\bar{u}(X_j))^{-1} \leqq nN . \tag{1.14}$$

By three applications of (12) and an application of (14), and by weakening the resulted bound we obtain

$$(1.15)\qquad n(2N+6)^{-1}\sum_{j=1}^{n} P_j(\text{rhs}\,(11))$$
$$\leqq (n-1)\bigvee_{y\geqq 0}\sum_{j=1}^{n} P_j\{(\bar{u}(X_j))^{-1}\boldsymbol{P}_x|\bar{u}_j(X_j-y)-\bar{\hat{u}}_j(X_j-y)|\}$$
$$+2mNnh^{-1}\,.$$

Since by the triangular inequality and Hölder's inequality,

$$(1.16)\qquad \boldsymbol{P}_x|\bar{u}_j(y)-\bar{\hat{u}}_j(y)|\leqq|\bar{u}_j(y)-\boldsymbol{P}_x\bar{\hat{u}}_j(y)|+\sigma_n(y)$$

where $\sigma_n^2(y)$=variance of $\bar{\hat{u}}_j(y)$, to get an upper bound of the first term of rhs (15) we shall obtain bounds for $\bigvee_y \sigma_n(y)$ and $\bigvee_{y\geqq 0}\sum_{j=1}^{n} P_j$((first term of rhs (16) at $X_j-y)/\bar{u}(X_j)$).

LEMMA 2.3.

$$\bigvee_y \sigma_n(y)\leqq m((n-1)h)^{-1/2}\,.$$

PROOF. By the definition of $\sigma_n^2$

$$(1.17)\qquad ((n-1)h)^2\sigma_n^2(y)\leqq \sum_{(j\neq)i=1}^{n}\boldsymbol{P}_x(\hat{u}_i(y))^2\leqq(n-1)\int_y^{y+h}\bar{u}_j(z)/f(z)dz$$

which is no more than $m^2(n-1)h$ because of (13).

LEMMA 2.4. *For all* $y\geqq 0$,

$$(1.18)\qquad \sum_{j=1}^{n} P_j(|\bar{u}_j(X_j-y)-\boldsymbol{P}_x\bar{u}_j(X_j-y)|/\bar{u}(X_j))\leqq nhm\,.$$

PROOF. Since $\boldsymbol{P}_x\bar{\hat{u}}_j(z)=h^{-1}\int_z^{z+h}\bar{u}_j(t)dt=\int_0^1\bar{u}_j(z+hs)ds$, lhs (18) for every $y\geqq 0$ equals to

$$\sum_{j=1}^{n}\int\left|\int_0^1 \bar{u}_j]_{z-y}^{z-y+hs}ds\right|\cdot p_j(z)/\bar{u}(z)dz$$

which is no more than

$$\sum_{j=1}^{n}\int_c^{d+1}\int_0^1(n-1)^{-1}\sum_{(j\neq)i=1}^{n}q(\theta_i)([\theta_i-hs\leqq z-y<\theta_i]+[\theta_i-hs\leqq z-y-1<\theta_i])ds$$
$$\cdot p_j(z)/\bar{u}(z)dz\,.$$

Thus, interchanging integrations and also averages over respective $j$ and $i$ leads to

$$(1.19)\qquad \text{lhs}\,(18)\leqq\sum_{i=1}^{n}q(\theta_i)\int_0^1\int_c^{d+1}([\theta_i-hs\leqq z-y<\theta_i]+[\theta_i-hs\leqq z-y-1<\theta_i])$$
$$\cdot(n-1)^{-1}\sum_{(i\neq)j=1}^{n}p_j(z)/\bar{u}(z)dzds\,.$$

Since $(n-1)^{-1}\sum_{(i\neq)j=1}^{n} p_j(z)/\bar{u}(z)\leqq f(z)\leqq 1$, by a simple computation and $q(\cdot)\leqq m$

$$\text{rhs (19)}\leqq 2h\sum_{i=1}^{n} q(\theta_i)\int_0^1 sds=h\sum_{i=1}^{n} q(\theta_i)\leqq nhm .$$

We now go back to (15). By Lemmas 2.4 and 2.3 together with an application of Lemma 2.2 we obtain in view of (16) that

$$\text{rhs (15)}\leqq mNn^{3/2}h^{-1/2}+(n-1)nhm+2mNnh^{-1} .$$

Therefore, in view of (9), we finally obtain

THEOREM 1.1. *For all* $\boldsymbol{\theta}\in[c, d]^n$,

$$|D(\boldsymbol{\theta}, \boldsymbol{\phi})^*|\leqq 2(2N+6)\{3mN(nh)^{-1/2}+mh\} .$$

*Remark.* In Chapter III of Nogami [6], two one-stage procedures; one (denoted by $\boldsymbol{\theta}_T$) for $\mathscr{P}^*(f)$ under Lipshitz condition for $1/f$ and the other (denoted by $\boldsymbol{\phi}$) for $\mathscr{P}^*(1)$, both with a rate 1/4 are exhibited as a special case ($k=1$) of the $k$-extended problem. (From the structure of construction $\boldsymbol{\phi}$ cannot be extended to $\mathscr{P}^*(f)$.) In Chapter II of Nogami [6] Theorem 3 (Theorem 2.1 in Section 2 of this paper) shows that when $\boldsymbol{\theta}=\mathbf{0}$ and $f\equiv 1$ (note that in this case $\boldsymbol{\theta}_T$ and $\boldsymbol{\phi}^*$ are the same estimate for the zero sequence $\mathbf{0}$), $\boldsymbol{\theta}_T$ with $h^{-1}n^{-1/4}=O(1)$ has exact order $O(h^2)$ of convergence, and Theorems 4 and 5 there give a lower bound and an upper bound for $D(\mathbf{0}, \boldsymbol{\phi})$ at $f\equiv 1$, respectively. In this section we assume no Lipshitz condition for $1/f$ and from above Theorem 1.1 we can see that for $\boldsymbol{\phi}^*$ with a choice of $h=n^{-1/3}$ (up to constants) $|D(\boldsymbol{\theta}, \boldsymbol{\phi}^*)|=O(n^{-1/3})$, uniformly in $\boldsymbol{\theta}\in[c, d]^n$. Furthermore, from Theorem 2.1 in the next section we shall see that for this choice of $h$ $\boldsymbol{\phi}^*$ has the best lower bound $n^{-2/3}$ for $D(\mathbf{0}, \boldsymbol{\phi}^*)$ at $f\equiv 1$ and this shows that $D(\boldsymbol{\theta}, \boldsymbol{\phi}^*)$ converges to zero at a rate no faster than $n^{-2/3}$.

## 2. Lower bounds of the modified regret $D(\mathbf{0}, \boldsymbol{\phi}^*)$ when $f\equiv 1$

In this section we consider the uniform distribution $P=U[0, 1)$ over the interval $[0, 1)$ as the underlined family of distributions. Lemmas 2.1 through 2.4 will be furnished to prove forthcoming Theorem 2.1 which gives us lower bounds of $D(\mathbf{0}, \boldsymbol{\phi}^*)$. Theorem 2.2 is a derivation from Theorem 2.1 and somewhat Section 1 and will be stated without proof.

Let $X_1, \cdots, X_n$ be i.i.d. random variables with the common distribution $P=U[0, 1)$. Let $\boldsymbol{X}=(X_1, \cdots, X_{n+1})$. Here we consider $\boldsymbol{\phi}^*(\boldsymbol{X})=(\phi^*_{1,n+1}, \cdots, \phi^*_{n+1,n+1})$. Since $\phi^*_{1,n+1}, \cdots, \phi^*_{n+1,n+1}$ are identically distributed and since for all $j$, $\theta_{j,n+1}=0$, abbreviating $\phi^*_{n+1,n+1}$ to $\phi^*$ we see in view of (0.3) that the modified regret of $\boldsymbol{\phi}^*$ at $\boldsymbol{\theta}=0$ is given by

$$D(\mathbf{0}, \boldsymbol{\phi}^*)=\boldsymbol{P}\boldsymbol{\phi}^{*2} \tag{2.1}$$

where $\boldsymbol{P}=P_1\times\cdots\times P_{n+1}$.

For fixed $x\doteq X_{n+1}$, $\phi^*$ is written as

$$\phi^*=(x'\vee\varphi)\wedge x \tag{2.2}$$

where with $\bar{\hat{u}}(y)=(n+1)^{-1}\sum_{i=1}^{n+1}\hat{u}_i(y)$,

$$\varphi=x-\int_0^1\sum_{r=0}^{\infty}\bar{\hat{u}}(\cdot-r)]_{x'}^{x'+t}/\bar{\hat{u}}(x)dt\,. \tag{2.3}$$

We shall exhibit an explicit form of $\varphi$ in a.e. $\boldsymbol{P}_x$-sense in the following:

LEMMA 2.1. *For every* $x\in[0, 1)$,

$$\begin{aligned}\varphi=\Big\{&\sum_{j=1}^{n}(X_j-h)[x<X_j\leqq x+h]-h\sum_{j=1}^{n}[0\leqq X_j\leqq x]-h\\&+\sum_{j=1}^{n}[0\leqq X_j\leqq x'+h]\Big\}\Big/\sum_{j=1}^{n}[x<X_j\leqq x+h]\qquad \text{a.e. } \boldsymbol{P}_x\,.\end{aligned} \tag{2.4}$$

PROOF. Fix $j$ and note that as a function of $t\in[0, 1]$, $\sum_{r=0}^{\infty}[X_j-x'+r-h\leqq t<X_j-x'+r]$ is equal to zero, is equal to its first term, or is equal to the sum of its first two terms according to whether $1<X_j-x'-h$, $X_j-x'-h\leqq 1<X_j-x'$ or $X_j-x'\leqq 1$. Integrating over $t\in[0, 1]$ for each case gives

$$\begin{aligned}&\int_0^1\sum_{r=0}^{\infty}[X_j-x'+r-h\leqq t<X_j-x'+r]dt\\&\quad=(x+h-X_j)[x<X_j\leqq x+h]+h[X_j\leqq x]\,.\end{aligned}$$

Hence, it follows

$$\begin{aligned}&((n+1)h)\int_0^1\sum_{r=0}^{\infty}\bar{\hat{u}}(\cdot-r)]_{x'}^{x'+t}dt\\&\quad=\sum_{j=1}^{n+1}(x+h-X_j)[x<X_j\leqq x+h]+h\sum_{j=1}^{n+1}[X_j\leqq x]\\&\quad\quad-\sum_{j=1}^{n+1}\sum_{r=0}^{\infty}[x'-r<X_j\leqq x'-r+h]\,.\end{aligned} \tag{2.5}$$

But since $[x<x\leqq x+h]=0$, $[x\leqq x]=1$, $\sum_{r=0}^{\infty}[x'-r<x\leqq x'-r+h]=0$ and a.e. $\boldsymbol{P}_x$, $\sum_{r=1}^{\infty}[x'-r<X_j\leqq x'-r+h]=0$, we have

$$\text{rhs}\,(5)=x\sum_{j=1}^{n}[x<X_j\leqq x+h]-\sum_{j=1}^{n}(X_j-h)[x<X_j\leqq x+h]$$

$$+h\sum_{j=1}^{n}[X_j\leqq x]+h-\sum_{j=1}^{n}[x'<X_j\leqq x'+h]\,, \qquad \text{a.e. } \boldsymbol{P}_x\,.$$

On the other hand, since $[x<x\leqq x+h]=0$,

$$((n+1)h)\bar{\hat{u}}(x)=\sum_{j=1}^{n}[x<X_j\leqq x+h]\,.$$

Applying these to the definition of $\varphi$, we get the asserted expression for $\varphi$.

In this section we need only to deal with $\varphi$ for $x<1-h$, where the term $\sum_{j=1}^{n}[0\leqq X_j\leqq x'+h]$ (cf. rhs (4)) vanishes. We also recognize that for $x<1-h$, $\boldsymbol{P}_x[\varphi>x]=0$. Hence, $\phi^*$ has the following simpler form:

$$\phi^*=\begin{cases} x'\vee\varphi & \text{for } x\in[0,1-h) \\ (x'\vee\varphi)\wedge x\,, & \text{for } x\in[1-h,1)\,. \end{cases} \tag{2.6}$$

Now, we let

$$J=[\varphi\geqq x',\ x<1-h] \tag{2.7}$$

and recognize by (1) and the definition of $\phi^*$ that

$$D(0,\boldsymbol{\phi}^*)\geqq \boldsymbol{P}(\varphi^2 J)\,. \tag{2.8}$$

Let $\xrightarrow{\mathcal{D}}$ denote convergence in distribution. Also, $N(c,d)$ denotes the normal distribution with mean $c$ and variance $d$. To get lower bounds for $D(0,\boldsymbol{\phi}^*)$ (Theorem 2.1) we use the relation (8) and the fact that for fixed $x$, $h^{-1}\varphi J\xrightarrow{\mathcal{D}}-2^{-1}$ and $S_n\doteq(\sqrt{nh}\,\varphi+2^{-1}\sqrt{nh^3})J\xrightarrow{\mathcal{D}}N(0,x^2)$. We then apply a convergence theorem (cf. Loéve [5] 11.4, A(i)):

$$\text{If } U_n\xrightarrow{\mathcal{D}}U\,,\ \text{then } \varliminf \mathrm{E}\,U_n^2\geqq \mathrm{E}\,U^2\,, \tag{2.9}$$

where E means expectation, and Theorem A in Appendix. We shall first prepare Lemmas 2.2, 2.3 and 2.4 to prove the above two convergences in distribution for the proof of forthcoming Theorem 2.1.

Let $u=\sum_{j=1}^{n}[0\leqq X_j\leqq x]$, $v=\sum_{j=1}^{n}[x<X_j\leqq x+h]$ and $w=\sum_{j=1}^{n}X_j[x<X_j\leqq x+h]$. We also define

$$X=(w-hv-xv-h)/(hv)\,,$$

$$Y=(u-nx)/\sqrt{nx(1-x)} \qquad \text{and}$$

$$Z=(v-nh)/\sqrt{nh}\,.$$

Then, on the set $J$, $\varphi$ of the form (4) is alternatively written as

(2.10) $$\varphi=hX+\frac{x(nh)^{-1/2}Z}{1+(nh)^{-1/2}Z}-\frac{\sqrt{x(1-x)}\,n^{-1/2}Y}{1+(nh)^{-1/2}Z}.$$

LEMMA 2.2. *Given $x\in(0,1)$, if $h$ is a function of $n$ such that $nh\to\infty$ and $h\to 0$, then*

$$(Y,Z)\xrightarrow{\mathcal{D}}N(\underline{0},I)$$

*where $\underline{0}$ is the zero vector in $R^2$ and $I$ is $2\times 2$ identity matrix.*

PROOF. For each $x\in(0,1)$ we restrict to $n$ such that $x<1-h$. Pick $t$ and $s$ arbitrary, and let

$$V_j=n^{-1/2}\{s(x(1-x))^{-1/2}([0\leqq X_j\leqq x]-x)+th^{-1/2}([x<X_j\leqq x+h]-h)\},$$

for $j=1,2,\cdots,n$. Then, it is not hard to see that

$$\sum_{j=1}^{n}V_j=sY+tZ.$$

Since the $V_j$ are i.i.d., the characteristic function $K$ of $(Y,Z)$ at a point $(s,t)\in R^2$ is given by

(2.11) $$K(s,t)=(J(1))^n$$

where $J$ is the characteristic function of $V\doteq V_1$.

Since by XV (6.8) (Feller, [2]), for any complex numbers such that $|\alpha|\leqq 1$ and $|\beta|\leqq 1$,

$$|\alpha^n-\beta^n|\leqq n|\alpha-\beta|,$$

(2.12) $$\left|(J(1))^n-\exp\left(-\frac{1}{2}(s^2+t^2)\right)\right|\leqq n\left|J(1)-\exp\left(-\frac{1}{2n}(s^2+t^2)\right)\right|.$$

By the triangular inequality and by using $|1-y-e^{-y}|=O(y^2)$ as $y\to 0$,

(2.13) $$\text{rhs}\,(12)\leqq n\left|J(1)-1+\frac{s^2+t^2}{2n}\right|+O(n^{-1}).$$

Now, from the Taylor development of characteristic functions by XV (4.14) (Feller, [2]) and from the fact that $J(0)=1$, $J'(0)=i\boldsymbol{P}_xV=0$ and $J''(0)=-\boldsymbol{P}_xV^2$, it follows that

$$\left|J(1)-1+\frac{1}{2}\boldsymbol{P}_xV^2\right|\leqq\frac{1}{6}\boldsymbol{P}_x|V|^3.$$

Now, we verify that

$$\boldsymbol{P}_xV^2=n^{-1}\{(s^2+t^2)-t^2h-2stx(\sqrt{x}\,(1-x)^{-1/2}+\sqrt{1-x}\,x^{-1/2})\sqrt{h}\}$$

and

$$\boldsymbol{P}_x|V|^3=n^{-3/2}\{|s(x^{-1}-1)^{1/2}-th^{1/2}|^3x+|t(1-h)h^{-1/2}-sx^{1/2}(1-x)^{-1/2}|^3h$$
$$+|sx^{1/2}(1-x)^{-1/2}+th^{1/2}|^3(1-x-h)\}\,.$$

Hence,

$$0\leqq n^{-1}(s^2+t^2)-\boldsymbol{P}_xV^2\leqq O(n^{-1}h^{1/2})$$

and

$$\boldsymbol{P}_x|V|^3=O(n^{-3/2}h^{-1/2})\,.$$

Hence, applying the triangular inequality leads to

$$\left|J(1)-1+\frac{s^2+t^2}{2n}\right|=O(n^{-1}h^{1/2}+n^{-3/2}h^{-1/2})\,.$$

Thus, in view of (13), (12) and (11),

$$\left|K(t,s)-\exp\left(-\frac{s^2+t^2}{2}\right)\right|=O(h^{1/2}+n^{-1/2}h^{-1/2}+n^{-1})\,.$$

To get the conclusion we invoke the continuity theorem (cf. e.g. Breiman [1], Theorem 11.6).

We shall next prove $X\xrightarrow{P}-2^{-1}$ where $\xrightarrow{P}0$ means convergence in probability $\boldsymbol{P}_x$ for given $x$.

LEMMA 2.3. *Under the same assumption as Lemma* 2.2,

$$X\xrightarrow{P}-2^{-1}\,.$$

PROOF. For given $x\in(0,1)$, we restrict to $n$ such that $x<1-h$. Then, $X$ is written as

$$X=\left(C\Big/\left(\frac{v}{nh}\right)\right)-v^{-1} \tag{2.14}$$

where $C=(nh)^{-1}\sum_{j=1}^{n}U_j$, where $U_j=h^{-1}(X_j-x-h)I_j$ with $I_j=[x<X_j\leqq x+h]$.

Since $v$ has the binomial distribution with parameters $n$ and $h$,

$$\frac{v}{nh}\xrightarrow{P}1 \text{ as } nh\to\infty \text{ and } h\to 0\,. \tag{2.15}$$

By simple computations,

$$\mathrm{E}\,U=-\frac{h}{2}$$

and

$$\operatorname{Var}(U)=\frac{h}{12}+\frac{h(1-h)}{4}.$$

Thus, $\mathrm{E}\,C=h^{-1}\,\mathrm{E}\,U=-2^{-1}$ and $\operatorname{Var}(C)=(nh^2)^{-1}\operatorname{Var}(U)=(12^{-1}+(1-h)/4)/(nh)$. Therefore, by the Chebychev inequality,

$$C\xrightarrow{P}-2^{-1}\ \text{as}\ nh\to\infty\ \text{and}\ h\to 0\,. \tag{2.16}$$

Applying (16), (15), (14) and Slutsky's theorem completes the proof of Lemma 2.3.

Besides the above two lemmas we shall show that $\boldsymbol{P}_x[\varphi\leqq x']$ vanishes when $nh\to\infty$ and $h\to 0$.

LEMMA 2.4. *Under the same assumption as Lemma 2.2,*

$$\boldsymbol{P}_x[\varphi\leqq x']\to 0 \qquad \textit{for fixed } x\,.$$

PROOF. We restrict to $n$ such that $x<1-h$. Let $W_j=h[0\leqq X_j\leqq x]-(X_j-h-x')[x<X_j\leqq x+h]$ for $j=1, 2, \cdots, n$. Then, by the representation (4) of $\varphi$, $[\varphi\leqq x']=[\bar{W}\geqq -n^{-1}h]$ where $\bar{W}$ is the average of i.i.d. $W_j$'s. Since $\boldsymbol{P}_x W_1=h(2^{-1}h+x')$,

$$\boldsymbol{P}_x[\varphi\leqq x']=\boldsymbol{P}_x[\bar{W}-\boldsymbol{P}_x\bar{W}\geqq(1-x-n^{-1}-2^{-1}h)h]\,. \tag{2.17}$$

But, $\operatorname{Var}(\bar{W})=n^{-1}\operatorname{Var}(W_1)=hn^{-1}\left\{1-(1-x)(2-x)h+\left(\frac{4}{3}-x\right)h^2-4^{-1}h^3\right\}\leqq\left(\frac{7}{3}\right)hn^{-1}$. Hence, by the Chebychev inequality and for large $n$

$$\text{rhs}\,(17)\leqq(7/3)h^{-1}n^{-1}(1-x-n^{-1}-2^{-1}h)^{-2}$$

which tends to zero when $nh\to\infty$ and $h\to 0$.

We are now ready to prove

THEOREM 2.1. (i) *If $h$ is a function of $n$ such that $nh^3\to\infty$ and $h\to 0$, then for any $\frac{1}{4}>\varepsilon>0$, there exists $N<+\infty$ so that for all $n\geqq N$*

$$D(0,\boldsymbol{\phi}^*)>\left(\frac{1}{4}-\varepsilon\right)h^2\,.$$

(ii) *If $h$ is a function of $n$ such that $nh\to\infty$, $h\to 0$ and $nh^3=O(1)$, then for any $\frac{1}{3}>\varepsilon>0$, there exists $N<+\infty$ so that for all $n\geqq N$*

$$D(0,\boldsymbol{\phi}^*)>\left(\frac{1}{3}-\varepsilon\right)\frac{1}{nh}\,.$$

PROOF. (i) Since $nh^3 \to \infty$ and $h \to 0$ implies $nh \to \infty$ and $h \to 0$, we have by Lemmas 2.2, 2.3 and 2.4 that given $x \in (0, 1)$,

$$(2.18) \qquad (Y, Z) \xrightarrow{\mathcal{D}} N(\underline{0}, I), \quad X \xrightarrow{P} -\frac{1}{2}, \quad [\varphi \geqq x'] \xrightarrow{P} 1.$$

Hence, in view of (10) it follows from Slutsky's theorem that if $x \in (0, 1)$, then $h^{-1}\varphi J \xrightarrow{\mathcal{D}} -2^{-1}$ (see (7) for the definition of $J$). By a convergence theorem (9), we have

$$(2.19) \qquad \underline{\lim}\, \boldsymbol{P}_x(h^{-2}\varphi^2 J) \geqq \frac{1}{4}[0<x<1],$$

and hence by Fatou's theorem applied to the lhs below

$$\underline{\lim}\, \mathrm{P}\, \boldsymbol{P}_x(h^{-2}\varphi^2 J) \geqq \mathrm{P}\,(\text{lhs}\,(19)) \geqq \frac{1}{4}.$$

Thus, by (8) we get that

$$\underline{\lim}\, h^{-2} D(\mathbf{0}, \boldsymbol{\phi}^*) \geqq \frac{1}{4}.$$

(i) follows because of the definition of lim inf.

To prove (ii) we first recognize that for this choice of $h$, (18) still holds. Let $S_n = \{\sqrt{nh}\,\varphi + 2^{-1}\sqrt{nh^3}\}J$. Then, in view of (10) it follows from Slutsky's theorem that if $x \in (0, 1)$, then

$$S_n \xrightarrow{\mathcal{D}} N(0, x^2).$$

Since $\boldsymbol{P}_x\{(nh)\varphi^2 J\} = \boldsymbol{P}_x(S_n - 2^{-1}\sqrt{nh^3}\,J)^2 \geqq \mathrm{Var}\,(S_n)$, applying Theorem A in Appendix to the rhs leads to

$$(2.20) \qquad \underline{\lim}\, \boldsymbol{P}_x\{(nh)\varphi^2 J\} \geqq x^2[0<x<1].$$

Thus, by Fatou's lemma applied to the lhs below

$$\underline{\lim}\, \mathrm{P}\, \boldsymbol{P}_x(nh\varphi^2 J) \geqq \mathrm{P}\,(\text{lhs}\,(20)) \geqq \int_0^1 y^2 dy = \frac{1}{3}.$$

Therefore by (8) we get that $\underline{\lim}\,(nh)D(\mathbf{0}, \boldsymbol{\phi}) \geqq \frac{1}{3}$ and the definition of lim inf leads to (ii).

Theorem 2.1 (i) implies that at any parameter sequence $(\theta_1, \theta_2, \cdots)$ where $\theta_1 = \theta_2 = \cdots$, $\boldsymbol{\phi}^*$ with the choice $h = n^{-1/3}$ has modified regret converging to zero at a rate no faster than $n^{-2/3}$.

*Remark.* By usage of the method obtaining Theorem 1.1 and the result of Theorem 2.1 we can verify the following:

THEOREM 2.2. (i) *If $h$ is a function of $n$ such that $nh^3 \to \infty$ and $h \to 0$, then there exists a positive constant $b_1$ such that for sufficiently large $n$,*

$$b_1^{-1}h^2 \leqq D(0, \boldsymbol{\phi}^*) \leqq b_1 h^2 .$$

(ii) *If $h$ is a function of $n$ such that $nh \to \infty$, $h \to 0$ and $nh^3 = O(1)$, then there exists a positive constant $b_2$ such that for sufficiently large $n$,*

$$b_2^{-1}(nh)^{-1} \leqq D(0, \boldsymbol{\phi}^*) \leqq b_2(nh)^{-1} .$$

From this theorem we can see that if $\boldsymbol{\phi}^*$ is defined by (1.7) with $h$ such that $nh^3 = b_0$, then there exists a positive constant $b_3$ so that for sufficiently large $n$, $b_3^{-1}n^{-2/3} \leqq D(0, \boldsymbol{\phi}^*) \leqq b_3 n^{-2/3}$. From this fact we may expect existence of $\boldsymbol{\phi}^*$ where $D(\boldsymbol{\theta}, \boldsymbol{\phi}^*)$ is of the best exact order $n^{-2/3}$, uniformly in $\boldsymbol{\theta} \in \Omega^n$.

## Appendix

The following theorem (A Fatou theorem for variances) is used in Section 2.

THEOREM A. *If $\{U_n\}$ is a sequence of random variables converging in distribution to a random variable $U$, then*

$$\underline{\lim} \operatorname{Var}(U_n) \geqq \operatorname{Var}(U) .$$

PROOF. It suffices to show that for $\{U_n\}$ such that $\operatorname{Var}(U_n) \to$ finite.

With $\mu_n = \mathrm{E}\, U_n$ and $\sigma_n^2 = \operatorname{Var} U_n$, the Chebychev inequality gives $\mathrm{P}\,[|U_n - \mu_n| < \sqrt{2}\,\sigma_n] \geqq 1/2$ while tightness provides a finite $b$ independent of $n$ for which $\mathrm{P}\,[|U_n| \leqq b] > 1/2$. The nonemptyness of the intersection of these events shows $|\mu_n| < b + \sqrt{2}\,\sigma_n$ so that $\{\mu_n\}$ is bounded.

Letting $\{\mu_m\}$ be a convergent subsequence with limit $\mu_\infty$, $U_m - \mu_m \xrightarrow{\mathscr{D}} U - \mu_\infty$ and hence (cf. Loéve [5] 11.4, A(i))

$$\lim \operatorname{Var}(U_n) = \lim \mathrm{E}\,(U_m - \mu_m)^2 \geqq \mathrm{E}\,(U - \mu_\infty)^2 \geqq \operatorname{Var} U .$$

## Acknowledgement

The author devotes sincere thanks to Professor James F. Hannan for his excellent guidance and encouragement during the research in

author's Ph.D. thesis [6].

DAITO BUNKA UNIVERSITY*

## REFERENCES

[1] Breiman, L. (1968). *Probability*, Addison-Wesley.

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

[3] Ferguson, T. S. (1967). *Mathematical Statistics and a Decision Theoretic Approach*, Academic Press, New York.

[4] Fox, R. (1968). Contribution to compound decision theory and empirical Bayes squared-error loss estimation, RM-214, Department of Statistics and Probability, Michigan State University.

[5] Loéve, Michel (1963). *Probability Theory* (*3rd* ed.), Van Nostrand, Princeton.

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

[7] Nogami, Y. (1978). The set-compound one-stage estimation in the nonregular family of distributions over the interval $(0, \theta)$, *Ann. Inst. Statist. Math.*, **30**, A, 35-43.

[8] Nogami, Y. (1979). The $k$-extended set-compound estimation problem in a nonregular family of distributions over $[\theta, \theta+1)$, *Ann. Inst. Statist. Math.*, **31**, A, 169-176.

[9] Singh, Radhey S. (1974). Estimation of derivatives of average of $\mu$-densities and sequence-compound estimation in exponential families, RM-318, Department of Statistics and Probability, Michigan State University.

[10] Wahba, Grace (1975). Optimal convergence properties of variable knot, kernel, and orthogonal series methods for density estimation, *Ann. Statist.*, **3**, 15-29.

* Now at The University of Tsukuba.