

THE k -EXTENDED SET-COMPOUND ESTIMATION PROBLEM IN
 A NONREGULAR FAMILY OF DISTRIBUTIONS OVER $[\theta, \theta+1)$

YOSHIKO NOGAMI

(Received Aug. 17, 1976; revised Nov. 28, 1978)

1. Introduction

Swain [8] has investigated, in his Ph. D. thesis, the use of more stringent standards in the compound decision problem and has called the resulting version the extended compound decision problem. Gilliland and Hannan [2] have developed more details for the nature of these standards.

The component problem in the compound decision problem here is the squared-error loss estimation (SELE) problem.

Let $k=1$. The component problem is the estimation of θ given $X \sim P_\theta \in \mathcal{P}$ where \mathcal{P} is the family of probability measures. Let $\mathbf{X}_n = (X_1, \dots, X_n)$ be n independent random variables with each $X_j \sim P_j \doteq P_\theta, \theta \in \mathcal{P}$ where \doteq is the defining property. Note that $k \leq n$. The set-compound problem is the estimation of θ_n according to the decision procedure $\mathbf{t}_n = (t_1, \dots, t_n)$ with each t_j depending on all n observations \mathbf{X}_n . With $L(\theta, a)$ denoting the loss the risk here is $R(\theta_n, \mathbf{t}_n) = \int n^{-1} \sum_{j=1}^n L(\theta_j, t_j(\mathbf{x}_n)) d(P_1 \times \dots \times P_n)(\mathbf{x}_n)$ and the modified regret for \mathbf{t}_n is $D(\theta_n, \mathbf{t}_n) = R(\theta_n, \mathbf{t}_n) - R(G_n)$ where G_n is the empirical distribution of $\theta_1, \theta_2, \dots, \theta_n$ and $R(G_n)$ is the Bayes envelope (the minimum Bayes risk) of the component problem at G_n which is the usual standard in the compound decision problem.

Let $k \geq 1$. Let $\mathbf{z}_j^k = (z_{j-k+1}, \dots, z_j), j = k, \dots, n$. Define by G_n^k the empirical distribution of the k -vectors $\theta_k^k, \theta_{k+1}^k, \dots, \theta_n^k$. Gilliland and Hannan [2] considered the following extended statistical game (the component problem): Player I picks $\theta_k = (\theta_1, \dots, \theta_k) \in \Omega^k$ where Ω is a parameter space of θ and Player II, after observing $\mathbf{X}_k \sim P_1 \times \dots \times P_k$, picks an action a according to some (nonrandomized) decision rule $t(\mathbf{X}_k)$. The risk Player II incurs is

$$R^k(\theta_k, t) = \int L(\theta_k, t(\mathbf{x}_k)) d(P_1 \times \dots \times P_k)(\mathbf{x}_k) .$$

The Bayes risk versus any prior G on Ω^k is

$$R^k(G, t) = \int R^k(\theta_k, t) dG(\theta_k).$$

Swain [8] used $R^k(G_n^k)$ as a standard for the k -extended compound problem. The modified regret for the k -extended set-compound problem is

$$(1.1) \quad D^k(\theta_n, \mathbf{t}_n) = \int (n-k+1)^{-1} \sum_{j=k}^n L(\theta_j, t_j(\mathbf{x}_n)) d(P_1 \times \cdots \times P_n)(\mathbf{x}_n) - R^k(G_n^k).$$

Swain [8] considered SELE problems in the discrete exponential and the normal families of distributions and obtained rates $O(n^{-1/4} \log^k n)$ and $o(1)$, respectively (all rates are uniformly in $\theta \in \mathcal{Q}^\infty$). Yu ([9], Part II) considered the same problems and obtained improved rates $O(n^{-1/2})$ (in Chapter 3) and $O(n^{-1/(k+4)})$ (in Chapter 4), respectively. The author ([4], Chapter III) has considered SELE problem in the nonregular (for the word "nonregular", see Ferguson ([1], p. 130)) family and showed the existence of the set-compound procedures with a rate $O(n^{-1/(2k+2)})$. In this paper she considers the same problem and introduces another procedure ϕ_n with an improved rate $O(n^{-1/3})$. To obtain a rate we apply the method used in Chapter 3 of Yu ([9], Part II). We note that this $O(n^{-1/3})$ is the same rate for the one-stage procedure ϕ^* in unextended ($k=1$) version of the same problem in Nogami [6].

Notational conventions. We abbreviate G_n^k and \mathbf{z}_n to G and \mathbf{z} , respectively. z' and z'_k abbreviate $z-1$ and (z_1-1, \dots, z_k-1) , respectively. Let $\mathbf{P} = P_1 \times \cdots \times P_n$ and $\bar{z} = (n-k+1)^{-1} \sum_{j=k}^n z_j$. For any function g , $g|_b^a$ means $g(a) - g(b)$ and let $g^k(\mathbf{z}_k) = \prod_{i=1}^k g(z_i)$. We often let Pg or $Pg(v)$ or $P(g)$ or $P(g(v))$ denote $\int g(v) dP(v)$. When we refer to (a. b) in Section a, we simply write (b). We denote the indicator function of a set A by A itself.

The Fubini theorem will often be applied without mentioning it.

2. Statement of the problem

Let ξ be Lebesgue measure and f be a measurable function such that for a given finite positive constant m ,

$$(2.1) \quad (0 <) m^{-1} \leq f \leq 1.$$

Define $q(\theta) = \left(\int_{\theta}^{\theta+1} f d\xi \right)^{-1}$ and then

$$(2.2) \quad 1 \leq q \leq m.$$

Letting $p_\theta = dP_\theta/d\xi$ we denote by $\mathcal{P}^*(f)$ the family of probability measures given by

$$\mathcal{P}^*(f) = \{P_\theta: p_\theta = q(\theta)[\theta, \theta + 1]f, \text{ for every } \theta \in \Omega\}$$

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

Let X_1, \dots, X_n be n independent random variables with each $X_j \sim P_j \in \mathcal{P}^*(f)$. For each $j = k, k + 1, \dots, n$, let $x \doteq X_j$, $y \doteq X_{j-1}^{k-1}$ and $\underline{x} \doteq (y, x) = X_j^k$. Let θ_G be the k -extended procedure, not counting the first $k-1$ coordinates, whose components are Bayes against G : $\theta_G = (\theta_{kn}, \theta_{k+1n}, \dots, \theta_{nn})$ with, each j ,

$$(2.3) \quad \theta_{jn} = \int_{x'+}^x \theta_j q^k(\theta_j^k) dG(\theta_j^k) / \int_{x'+}^x q^k(\theta_j^k) dG(\theta_j^k)$$

where the affix $+$ is intended to describe the integration as over the interval $(x', x]$. Hereafter we suppress the affix $+$. The Bayes envelope in the k -extended problem is of form

$$R^k(G) = (n - k + 1)^{-1} \sum_{j=k}^n P(\theta_{jn} - \theta_j)^2$$

and hence, in view of (1.1), the modified regret for any set-compound procedure t relative to the k -extended envelope is given by

$$(2.4) \quad D^k(\theta, t) = (n - k + 1)^{-1} \sum_{j=k}^n \{P(\theta_j - t_j(X))^2 - P(\theta_{jn} - \theta_j)^2\}.$$

Since for each j $X_j' < \theta_{jn} \leq X_j$, we have that when $X_j' < t_j(X) \leq X_j$,

$$(2.5) \quad (n - k + 1)2^{-1} |D^k(\theta, t)| \leq \sum_{j=k}^n P |t_j(X) - \theta_{jn}|.$$

In Section 3 we shall construct the set-compound procedure ϕ and prove in Section 4 the following theorem:

THEOREM. For all $\theta \in \Omega^\infty$,

$$|D^k(\theta, \phi)| \leq (8N + 24)m^k \{N^k(4k - 2 + k^{1/2})((n - k + 1)h^{-k})^{1/2} + 2^{-k+1}h^k\}$$

where $N = d + 1 - c$.

From above theorem we immediately get

COROLLARY. For ϕ with a choice of $h = n^{-1/(3k)}$,

$$|D^k(\theta, \phi)| = O(n^{-1/3}), \quad \text{uniformly in } \theta \in \Omega^\infty.$$

3. The decision procedure ϕ

In this section we shall use Lemma 1 below to get an alternative form of θ_{jn} and estimate θ_{jn} by ϕ_{jn} (see (8) below for definition) through this alternative form of θ_{jn} . This estimate $\phi=(\phi_{kn}, \phi_{k+1n}, \dots, \phi_{nn})$ for θ_G can also be an estimate for θ .

Let for an integer r R^r be the r -dimensional real line with $R=R^1$, $v=(w, v) \in R^{k-1} \times R$ and similarly $\theta=(w, \theta)=R^{k-1} \times R$. We denote by Q the measure with density $q^k(v)$ at v wrt G . In view of (2.3),

$$(3.1) \quad \theta_{jn} = \int_{x'}^x \theta dQ(\theta) / \int_{x'}^x dQ(\theta), \quad \text{for } j=k, k+1, \dots, n.$$

Following Lemma 1 is a generalization of Lemma 1.1 of Nogami [6] with $g=1$.

LEMMA 1. Let τ be a signed measure and $I=(v', v]$ be a cube with $\tau(I) \neq 0$. Let τ_v be the signed measure with density $I/\tau(I)$ wrt τ . Then,

$$(3.2) \quad \int s d\tau_v(s) = v - \int_0^1 \tau_v[s \leq v' + t] dt.$$

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

$$\int (v-s) d\tau_v(s) = \int \int_{s-v'}^1 dt d\tau_v(s) = \int_0^1 \tau_v[s \leq v' + t] dt.$$

Applying to rhs(1) Lemma 1 with τ the measure with density $(Q(x', x))^{-1}$ wrt Q gives us that

$$(3.3) \quad \theta_{jn} = x - \left(\int_0^1 Q((y', y) \times (x', x' + t]) dt / Q(x', x) \right).$$

For every $v \in R^k$, let $p_\theta(v) = p_w(w)p_\theta(v) = \left(\prod_{i=1}^{k-1} p_{w_i}(w_i) \right) p_\theta(v)$. We furthermore define the following: for $i=k, k+1, \dots, n$

$$(3.4) \quad u_i(v) = p_{\theta_i^k}(v) / f^k(v) = q^k(\theta_i^k) [v' < \theta_i^k \leq v].$$

By the definition of Q we can easily check

$$(3.5) \quad \bar{u}(v) = Q((w', w] \times (-\infty, \cdot])|_v.$$

By a telescopic series applied to v -coordinate,

$$(3.6) \quad Q((w', w] \times (-\infty, v]) = \sum_{r=0}^{\infty} \bar{u}(w, v-r)$$

where the summation wrt the nonnegative integer r involves at most

$d-c+2$ ($\doteq N+1$) terms. From an application of (5) to the denominator in (3) and two applications of (6) to the integrand of the numerator in (3), we get a final alternative form of θ_{j_n} :

$$(3.7) \quad \theta_{j_n} = x - \left(\int_0^1 \sum_{r=0}^{\infty} \bar{u}(y, \cdot) \Big|_{x'-r}^{x'-r+t} dt / \bar{u}(x) \right).$$

In view of (4) we estimate $\bar{u}(v)$ by $\bar{u}(v) = (n-k+1)^{-1} \sum_{i=k}^n \hat{u}_i(v)$ where for any $h > 0$

$$\hat{u}_i(v) = h^{-k} [v \leq X_i^k < v+h] / f^k(X_i^k)$$

with $\underline{1} = (1, 1)$, the k -dimensional 1-vector, h can depend on n and $h < 1$ for convenience. Thus, this, (7) and the fact that $x' < \theta_{j_n} \leq x$ suggest an estimate ϕ_{j_n} of θ_{j_n} given by

$$(3.8) \quad \phi_{j_n} = x - 0 \vee \left\{ \int_0^1 \sum_{r=0}^{\infty} \bar{u}(y, \cdot) \Big|_{x'-r}^{x'-r+t} dt / \bar{u}(x) \right\} \wedge 1.$$

Hence $\phi = (\phi_{1_n}, \dots, \phi_{n_n})$ is an estimate of $\theta_G(X)$ and thus of θ .

In the next section we shall prove Theorem in Section 2 which gives us an upper bound of the modified regret $D^k(\theta, \phi)$.

4. A proof of Theorem

Let $P_j^k = P_{j-k+1} \times \dots \times P_j$ and $\check{P}_{j,k} = P_1 \times \dots \times P_{j-k} \times P_{j+1} \times \dots \times P_n$. In view of (2.5)

$$(4.1) \quad (n-k+1)2^{-1} |D^k(\theta, \phi)| \leq \sum_{j=k}^n P_j^k (\check{P}_{j,k} |\phi_{j_n} - \theta_{j_n}|).$$

As we have done in Nogami [5], [6] applying Lemma A.2 of Singh [7] (see Corollary 1.1 in Nogami [6]) and weakening the resulted bound leads to

$$(4.2) \quad 2^{-1} \check{P}_{j,k} |\phi_{j_n} - \theta_{j_n}| \leq (\bar{u}(x))^{-1} \left\{ \sum_{r=1}^{N+1} \left\{ \int_0^1 \check{P}_{j,k} |\bar{u}(x-r+t) - \bar{u}(x-r+t)| dt + \check{P}_{j,k} |\bar{u}(x-r) - \bar{u}(x-r)| \right\} + 2\check{P}_{j,k} |\bar{u}(x) - \bar{u}(x)| \right\}$$

where $\underline{x-z} = (y, x-z)$. Before getting a bound of rhs (1) we shall introduce an unattainable estimate \bar{u}^* of \bar{u} as Yu [9] has done.

Fix $j \in \{k, k+1, \dots, n\}$ until (4). Let $A_j = \{j-k+1, \dots, j+k-1\}$ and $X_{i,k}^* = (X_{i-k+1}, \dots, X_{j-k}, X_{j-k+1}^*, \dots, X_i^*)$ for $i < j$; $= X_j^{*k}$ for $i = j$; $= (X_{i-k+1}^*, \dots, X_j^*, X_{j+1}, \dots, X_i)$ for $i > j$, where X_r^* is independently distributed according to P_r and independent of X_r . For any real $v \in R^k$,

define $\hat{u}_i^*(v)$ for $i=k, k+1, \dots, n$ by

$$(4.3) \quad \hat{u}_i^*(v) = \begin{cases} \hat{u}_i(v), & \text{for } i \notin A_j \\ (h^k f^k(X_{i,k}^*))^{-1} [v \leq X_{i,k}^* < v+h\mathbb{1}], & \text{for } i \in A_j. \end{cases}$$

Then, we can estimate $\bar{u}(v)$ by an unattainable estimate $\bar{u}^*(v)$.

Let E_j be the measure on $(X_j^*, (X_1, \dots, X_{j-k}, X_{j-1}, \dots, X_n))$. By two usages of the triangular inequality

$$\begin{aligned} \check{P}_{j,k} |\bar{u}(v) - \bar{u}^*(v)| &\leq E_j |\bar{u}(v) - \bar{u}^*(v)| + E_j |\bar{u}^*(v) - E_j \bar{u}^*(v)| \\ &\quad + |E_j \bar{u}^*(v) - \bar{u}(v)|. \end{aligned}$$

Since by Hölder's inequality $(E_j |\bar{u}^*(v) - E_j \bar{u}^*(v)|)^2 \leq \sigma_n^2(v)$ where $\sigma_n^2(v) =$ variance of $\bar{u}^*(v)$, and since

$$\begin{aligned} (n-k+1)h^k |\bar{u}(v) - \bar{u}^*(v)| &\leq \sum_{i=j-k+1}^{j+k-1} [v \leq X_i^k < v+h\mathbb{1}] / f^k(X_i^k) \\ &\leq 2m^k(2k-1), \end{aligned}$$

we obtain

$$(4.4) \quad \begin{aligned} \check{P}_{j,k} |\bar{u}(v) - \bar{u}^*(v)| &- (4k-2)m^k((n-k+1)h^k)^{-1} \\ &\leq \sigma_n(v) + |E_j \bar{u}^*(v) - \bar{u}(v)|. \end{aligned}$$

In view of (2) and (1),

$$(4.5) \quad \begin{aligned} (4N+12)^{-1}(\text{rhs (1)}) \\ &\leq \sum_{j=k}^n P_j^k \{ \{ (4k-2)m^k((n-k+1)h^k)^{-1} + \sup_v \sigma_n(v) \} / \bar{u}(x) \} \\ &\quad + \sup_z \sum_{j=k}^n P_j^k \{ |E_j \bar{u}^*(x-z) - \bar{u}(x-z)| / \bar{u}(x) \}. \end{aligned}$$

Hence, we shall obtain upper bounds for $\sum_{j=k}^n P_j^k (\bar{u}(x))^{-1}$, $\sup_v \sigma_n(v)$ and (the second term of rhs (5)) in Lemmas 2, 3 and 4, respectively.

LEMMA 2.

$$(4.6) \quad \sum_{j=k}^n P_j^k (\bar{u}(x))^{-1} \leq (n-k+1)N^k.$$

PROOF. Since by the definition of $\bar{u} \sum_{j=k}^n q^k(\theta_j^k) [\theta_j^k \leq x < \theta_j^k + \mathbb{1}] = (n-k+1)\bar{u}(x)$, it follows by k usages of $f \leq 1$ that $(n-k+1)^{-1}(\text{lhs (6)}) = \int \{v: \bar{u}(v) > 0\} f^k(v) dv \leq N^k$.

LEMMA 3. For every $v \in R^k$,

$$\sigma_n(v) \leq k^{1/2} m^k ((n-k+1)h^k)^{-1/2}.$$

PROOF. Let \sum' denote summation over $\lambda=0, 1, 2, \dots, k-1$ and \sum'' denote summation over integers r for which $k \leq \lambda + rk \leq n$. Then,

$$(4.7) \quad (n-k+1)^2 \sigma_n^2(v) = E_j \{ \sum' \sum'' (E_j \hat{u}_{i+rk}^*(v) - \hat{u}_{i+rk}^*(v)) \}^2.$$

By Minkowski inequality (cf. Loève [3], p. 156) applied to lhs (8) of the first inequality below,

$$(4.8) \quad \begin{aligned} \text{rhs (7)}^{1/2} &\leq \sum' E_j^{1/2} \{ \sum'' (E_j \hat{u}_{i+rk}^*(v) - \hat{u}_{i+rk}^*(v)) \}^2 \\ &\leq \sum' (\sum'' E_j \hat{u}_{i+rk}^{*2}(v))^{1/2}. \end{aligned}$$

Since for $i=k, k+1, \dots, n$

$$E_j \hat{u}_i^{*2}(v) = h^{-2k} \int_{\underline{v}}^{\underline{v}+h\underline{1}} (f^k(t))^{-1} q^k(\theta_i^k) [\theta_i^k \leq \underline{v} < \theta_i^k + \underline{1}] dt \leq m^{2k} h^{-k}$$

where the inequality follows by (2.1) and (2.2),

$$\text{rhs (8)} \leq k^{1/2} (n-k+1)^{1/2} m^k h^{-k/2}.$$

This, (8) and (7) gives the bound of Lemma 3.

LEMMA 4. For every $z \in R$,

$$(4.9) \quad \sum_{j=k}^n P_j^k \{ |E_j \bar{u}^*(\underline{x}-z) - \bar{u}(\underline{x}-z)| / \bar{u}(\underline{x}) \} \leq 2^{-k+1} m^k h^k (n-k+1).$$

PROOF. Since $E_j \bar{u}^*(\underline{x}-z) = h^{-k} \int_{\underline{v}}^{\underline{v}+h\underline{1}} \int_{\underline{x}-z}^{\underline{x}-z+h\underline{1}} \bar{u}(v) dv$, by a change of variables v to $\underline{t} = (v - \underline{x} - z) / h\underline{1}$

$$\begin{aligned} &|E_j \bar{u}^*(\underline{x}-z) - \bar{u}(\underline{x}-z)| \\ &\leq \left| \int_0^{\underline{1}} (\bar{u}(\underline{x}-z+h\underline{1}t) - \bar{u}(\underline{x}-z)) dt \right| \\ &\leq \int_0^{\underline{1}} (n-k+1)^{-1} \sum_{i=k}^n q^k(\theta_i^k) \{ [\theta_i^k - h\underline{1}t < \underline{x} < \theta_i^k] \\ &\quad + [\theta_i^k + \underline{1} - h\underline{1}t < \underline{x} < \theta_i^k + \underline{1}] \} dt. \end{aligned}$$

Thus, by k applications of $f \leq 1$ and $q \leq m$ to the lhs of the second inequality below

$$\begin{aligned} \text{lhs (9)} &\leq \sum_{i=k}^n q^k(\theta_i^k) \int_{\underline{t}=0}^{\underline{1}} \int_{\underline{v}=\underline{c}}^{\underline{d}+\underline{1}} \{ [\theta_i^k - h\underline{1}t < \underline{v} < \theta_i^k] + [\theta_i^k + \underline{1} - h\underline{1}t < \underline{v} < \theta_i^k + \underline{1}] \} \\ &\quad \cdot f^k(v) dv dt \\ &\leq 2m^k h^k (n-k+1) \int_0^{\underline{1}} \underline{t} dt \end{aligned}$$

which gives rhs (9).

We now go back to the inequality (5). Applying Lemmas 2, 3 and 4 leads to

$$\text{rhs (5)} \leq m^k N^k \{(4k-2)h^{-k} + k^{1/2}((n-k+1)h^{-k})^{1/2}\} + 2^{-k+1} m^k (n-k+1) h^k .$$

Therefore, in view of (1) and (5) and by weakening the bound we finally obtain asserted bound in Theorem in Section 2.

Acknowledgement

The author wishes to express her sincere appreciation to Professor James F. Hannan for his variable suggestions and comments on her Ph. D. thesis [4].

DAITO BUNKA UNIVERSITY

REFERENCES

- [1] Ferguson, T. S. (1967). *Mathematical Statistics and a Decision Theoretic Approach*, Academic Press, New York.
- [2] Gilliland, D. C. and Hannan, J. F. (1969). On an extended compound decision problem, *Ann. Math. Statist.*, **40**, 1536-1541.
- [3] Loève, Michel (1963). *Probability Theory* (3rd ed.), Van Nostrand, Princeton.
- [4] Nogami, Y. (1975). A nonregular squared-error loss set-compound estimation problem, RM-345, Department of Statistics and Probability, Michigan State University.
- [5] Nogami, Y. (1978a). 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.
- [6] Nogami, Y. (1978b). The set-compound one-stage estimation in the nonregular family of distributions over the interval $[\theta, \theta+1)$. (Submitted to *Ann. Inst. Statist. Math.*)
- [7] Singh, Radhey S. (1974). Estimation of derivatives of average of μ -densities and sequence-compound estimation in exponential families, RM-318, Department of Statistics and Probability, Michigan State University.
- [8] Swain, Donald D. (1965). Bounds and rates of convergence for the extended compound estimation problem in the sequence case, *Technical Report* No. 81, Department of Statistics, Stanford University.
- [9] Yu, Benito Ong (1971). Rates of convergence in empirical Bayes two-action and estimation problems and in extended sequence-compound estimation problems, RM-279, Department of Statistics and Probability, Michigan State University.