Ann. Inst. Statist. Math.
35 (1983), Part A, 175-183

SOME PROPERTIES OF THE RISK SET IN MULTIPLE DECISION PROBLEMS

MASAKATSU MURAKAMI

(Received Mar. 22, 1979; revised May 30, 1983)

Summary

Some properties of the risk set of a decision problem with n-action, m-sample and 2-parameter are considered. It is shown that the number of vertices of the risk set is equal to $mn-(t_1+t_2)$, and that the number of essentially nonrandomized decision rules (defined in Section 1) in the minimal complete class is equal to $m(n-1)+1-t_1$, where t_1 and t_2 are defined in Section 2. Also, a procedure is given for getting all nonrandomized decision rules in the minimal complete class.

1. Introduction

Let $L(\theta, a)$ be the loss incurred by an action a when the parameter value is θ. Let $f(x|\theta)$ be the probability distribution of a sample x when the parameter value is θ.

We consider the following situation (Decision problem A): let $\Theta=\{\theta_1, \theta_2\}$, $\mathcal{X}=\{x_1, \cdots, x_m\}$, and $\mathcal{A}=\{a_1, \cdots, a_n\}$, be the parameter, the sample and the action spaces, respectively. We assume

$$\begin{aligned} & f(x|\theta)>0\,, \qquad \text{for } x\in\mathcal{X}, \text{ and } \theta\in\Theta\,, \\ (1) \qquad & L(\theta_1, a_1)<L(\theta_1, a_2)<\cdots<L(\theta_1, a_n)\,, \quad \text{and} \\ & L(\theta_2, a_1)>L(\theta_2, a_2)>\cdots>L(\theta_2, a_n)\,. \end{aligned}$$

To avoid any reduction of the problem, we further assume that the action a_i with $1<i<n$ satisfies the condition

$$(2) \qquad \frac{L(\theta_1, a_i)-L(\theta_1, a_{i-1})}{L(\theta_1, a_{i+1})-L(\theta_1, a_i)}\cdot\frac{L(\theta_2, a_{i+1})-L(\theta_2, a_i)}{L(\theta_2, a_i)-L(\theta_2, a_{i-1})}<1\,.$$

Let D be the set of all nonrandomized decision rules, the mapping from the sample space $\mathcal{X}$ to the action space $\mathcal{A}$. Each $d\in D$ can be expressed in the coordinate form as

$$d=(a_{i_1}, a_{i_2}, \cdots, a_{i_m})$$

where

$$a_{i_k}=d(x_k)\,, \qquad k=1, 2, \cdots, m\,.$$

Let $\mathcal{D}$ be the set of all convex linear combinations δ of nonrandomized decision rules:

$$\delta=\sum_{j=1}^{t} \pi_j d_j$$

where $d_j \in D$, $\pi_j \geqq 0$ and $\sum_{j=1}^{t} \pi_j=1$. We call δ a randomized decision rule. The risks of $d \in D$ and $\delta \in \mathcal{D}$ are defined by

$$R(\theta, d)=\mathrm{E}_\theta\, L(\theta, d(x))=\sum_{k=1}^{m} L(\theta, d(x_k))f(x_k|\theta)$$

and,

$$R(\theta, \delta)=\sum_{j=1}^{t} \pi_j R(\theta, d_j)$$

respectively.

We say that δ is better than δ' if $R(\theta, \delta)\leqq R(\theta, \delta')$ for all $\theta \in \Theta$ with an exact inequality holding for at least one θ. Decision rules δ and δ' are said to be equivalent if $R(\theta, \delta)=R(\theta, \delta')$ for all $\theta \in \Theta$. A nonrandomized decision rule d is said to be *essentially nonrandomized* if no randomized decision rule is equivalent to d. A rule δ is said to be admissible if no rule is better than δ. A subclass $\mathcal{C}$ of $\mathcal{D}$ is said to be complete, if for any given rule δ not in $\mathcal{C}$, there exists a rule in $\mathcal{C}$ that is better than δ. A complete class $\mathcal{C}_1$ is said to be minimal complete if no proper subclass of $\mathcal{C}_1$ is complete.

In the following section, we investigate the properties of the risk set $\mathcal{S}$ of Problem A. In this decision problem, the risk set is a convex polygon and its vertices correspond to essentially nonrandomized decision rules. We shall give in Theorem 1, a procedure for obtaining all nonrandomized decision rules in $\mathcal{C}_1$. Secondly, using Theorem 1 we give the number N_1 of essentially nonrandomized decision rules in $\mathcal{C}_1$ (Theorem 2) and the number N_3 of vertices of the risk set $\mathcal{S}$ (Theorem 3).

2. The main result

We write $\Delta(d, d')$ to mean the slope of the line connecting the two risk points $(R(\theta_1, d), R(\theta_2, d))$ and $(R(\theta_1, d'), R(\theta_2, d'))$:

$$\Delta(d, d')=\frac{R(\theta_2, d')-R(\theta_2, d)}{R(\theta_1, d)-R(\theta_1, d')}\,.$$

The following lemma will be needed later.

LEMMA 1. *Let d^* be a nonrandomized admissible decision rule. Let $D^-(d^*)$ be the set of all nonrandomized decision rules d such that*

$$R(\theta_2, d) < R(\theta_2, d^*) . \tag{3}$$

If $D^-(d^*) \neq \phi$ and if a decision rule d^{**} satisfies

$$\Delta(d^*, d^{**}) = \operatorname*{Max}_{d \in D^-(d^*)} \{\Delta(d^*, d)\} , \tag{4}$$

then d^{**} is admissible.

PROOF. The assertion is a direct consequence of the very definition of d^{**}.

For a nonrandomized decision rule $d=(a_{i_1}, a_{i_2}, \cdots, a_{i_m})$, let $d'=(a_{i'_1}, \cdots, a_{i'_m}) \in D'(d)$ be the nonrandomized decision rule such that

$$i'_k = \begin{cases} i_k + 1 & \text{for only one } k \ (i'_k \leqq n-1) \text{ and} \\ i_k & \text{for other } k. \end{cases}$$

THEOREM 1. *In the Problem A suppose that a nonrandomized decision rule d^* is admissible, then the nonrandomized decision rule d^{**} such that*

$$\Delta(d^*, d^{**}) = \max_{d' \in D'(d^*)} \{\Delta(d^*, d')\}$$

is also admissible.

PROOF. We are to show that

$$\operatorname*{Max}_{d' \in D'(d^*)} \{\Delta(d^*, d')\} = \operatorname*{Max}_{d \in D^-(d^*)} \{\Delta(d^*, d)\} .$$

Putting $d^*(x_k) = a_{i_k}$ and $d(x_k) = a_{i_k + \alpha_k}$, $k=1, 2, \cdots, n$, we have for $d \in D^-(d^*)$,

$$\begin{aligned} \Delta(d^*, d) &= \frac{R(\theta_2, d) - R(\theta_2, d^*)}{R(\theta_1, d^*) - R(\theta_1, d)} \\ &= \frac{\sum_{k=1}^{m} \{L(\theta_2, d(x_k)) - L(\theta_2, d^*(x_k))\} f(x_k | \theta_2)}{\sum_{k=1}^{m} \{L(\theta_1, d^*(x_k)) - L(\theta_1, d(x_k))\} f(x_k | \theta_1)} \\ &= \frac{\sum_{k=1}^{m} \{L(\theta_2, a_{i_k+\alpha_k}) - L(\theta_2, a_{i_k})\} f(x_k | \theta_2)}{\sum_{k=1}^{m} \{L(\theta_1, a_{i_k}) - L(\theta_1, a_{i_k+\alpha_k})\} f(x_k | \theta_1)} . \end{aligned}$$

On the other hand if $\Delta(d^*, d)$ attains its maximum at $d = d^{**} \in D'(d^*)$ then we have for some k',

$$\Delta(d^*, d^{**}) = \frac{R(\theta_2, d^{**}) - R(\theta_2, d^*)}{R(\theta_1, d^*) - R(\theta_1, d^{**})}$$
$$= \frac{\{L(\theta_2, a_{i_{k'}+1}) - L(\theta_2, a_{i_{k'}})\} f(x_{k'}|\theta_2)}{\{L(\theta_1, a_{i_{k'}}) - L(\theta_1, a_{i_{k'}+1})\} f(x_{k'}|\theta_1)} .$$

But the condition (2) implies that

$$\frac{L(\theta_2, a_{i_{k'}+1}) - L(\theta_2, a_{i_{k'}})}{L(\theta_1, a_{i_{k'}}) - L(\theta_1, a_{i_{k'}+1})}$$

is less than or larger than

$$\frac{L(\theta_2, a_{i_{k'}+\alpha_{k'}}) - L(\theta_2, a_{i_{k'}})}{L(\theta_1, a_{i_{k'}}) - L(\theta_1, a_{i_{k'}+\alpha_{k'}})}$$

according as $\alpha_k < 0$ or $\alpha_{k'} > 1$. Then, we have

$$\Delta(d^*, d^{**}) \geqq \max_{d \in D^-(d^*)} \{\Delta(d^*, d)\} ,$$

by using the inequality* that if y_i, y'_i, z_i and z'_i are positive numbers such

$$\frac{z_i}{y_i} \geqq \frac{z_k}{y_k} \qquad \text{for } i=1, \cdots, n ,$$

$$\frac{z'_i}{y'_i} \leqq \frac{z_k}{y_k} \qquad \text{for } i=1, \cdots, m ,$$

and if $\sum_{i=1}^{n} y_i < \sum_{i=1}^{m} y'_i$, then

$$\frac{\sum_{i=1}^{n} z_i - \sum_{i=1}^{m} z'_i}{\sum_{i=1}^{n} y_i - \sum_{i=1}^{m} y'_i} \leqq \frac{z_k}{y_k} .$$

(Put $\quad y = \{L(\theta_1, a_i) - L(\theta_1, a_{i-\alpha})\} f(x|\theta_1)$,

---

*) Since

$$\frac{z_k}{y_k} \leqq \frac{\sum_{i=1}^{n} z_i}{\sum_{i=1}^{n} y_i} \quad \text{and} \quad \frac{z_k}{y_k} \geqq \frac{\sum_{i=1}^{m} z'_i}{\sum_{i=1}^{m} y'_i} ,$$

it follows that

$$\frac{\sum_{i=1}^{n} z_i - \sum_{i=1}^{m} z'_i}{\sum_{i=1}^{n} y_i - \sum_{i=1}^{m} y'_i} - \frac{z_k}{y_k} = \frac{Q}{\left(\sum_{i=1}^{n} y_i - \sum_{i=1}^{m} y'_i\right) y_k} < 0 ,$$

where

$$Q = y_k \sum_{i=1}^{n} z_i - z_k \sum_{i=1}^{n} y_i + z_k \sum_{i=1}^{m} y'_i - y_k \sum_{i=1}^{m} z'_i .$$

$$z=\{L(\theta_2, a_{i-\alpha})-L(\theta_2, a_i)\}f(x|\theta_2)\ ,$$

$$y'=\{L(\theta_1, a_i)-L(\theta_1, a_{i+\alpha})\}f(x|\theta_1) \quad \text{and}$$

$$z'=\{L(\theta_2, a_{i+\alpha})-L(\theta_2, a_i)\}f(x|\theta_2) \quad \text{where } \alpha>0\ .)$$

The desired equality follows from $D'(d^*)\subset D^-(d^*)$ and we conclude from Lemma 1 that d^{**} is admissible.

3. Some properties of the risk set

In this section we investigate some properties of the risk set using Theorem 1. A decision rule δ_0 is said to be *unfavorable* if there exists no decision rule $\delta\in\mathcal{D}$ such that

$$R(\theta, \delta_0)\leqq R(\theta, \delta) \qquad \text{for all } \theta\in\Theta$$

and

$$R(\theta, \delta_0)<R(\theta, \delta) \qquad \text{for at least one } \theta\in\Theta\ .$$

Let C_2 be the set of all unfavorable decision rules and let N_1, N_2 and N_3 be the numbers of essentially nonrandomized decision rules in C_1, C_2, and $\mathcal{S}$, respectively.

Write

$$V(i, j, k)=\frac{\{L(\theta_2, a_i)-L(\theta_2, a_j)\}f(x_k|\theta_2)}{\{L(\theta_1, a_j)-L(\theta_1, a_i)\}f(x_k|\theta_1)}\ .$$

THEOREM 2. *Let t_1 and t_2 be numbers of quardruplet (i, i', k, k') and doublet (k, k') which satisfy*

$$V(i, 1, k)=V(i', 1, k') \tag{5}$$

and

$$V(1, n, k)=V(1, n, k') \tag{6}$$

where $1\leqq i, i'\leqq n-1$ and $1\leqq k, k'\leqq m$.

In Problem A, the condition (2) *implies*

(a) $N_1=m(n-1)+1-t_1$

(b) $N_2=m+1-t_2$ *and*

(c) $N_3=mn-(t_1+t_2)$.

PROOF. (a) Since $\mathcal{S}$ is bounded from below and closed from below, the minimal complete class exists and it consists exactly of all the admissible rules (See [1], p. 56 and p. 69). We first show that if the condition

$$(7) \qquad \nabla(i, 1, k) \neq \nabla(i', 1, k')$$

for $1 \leqq i, i' \leqq n-1$ and $1 \leqq k, k' \leqq m$ is satisfied, then $N_1 = m(n-1)+1$.

Let d_0 be a rule for which $d_0(x_i) = a_1$ for all i, then it is easy to see that d_0 is admissible. We use the notation $\wedge$ (hat) to show a rule is admissible. By Theorem 1 any rule $d \in D'(\hat{d}_0)$ which satisfies

$$\Delta(\hat{d}_0, d) = \max_{d \in D'(d_0)} \{\Delta(\hat{d}_0, d)\}$$

is admissible. Starting from $\hat{d}_0$, we can find a sequence $\{\hat{d}_i\}$ $i=0, 1, \cdots, m(n-1)$ of admissible decision rules as follows. There exists exactly one rule d in $D'(\hat{d}_0)$ which satisfies (7). We denote it by $\hat{d}_1$. Similarly, if $\hat{d}_i$ is given, we can find an admissible rule $\hat{d}_{i+1}$ which satisfies

$$(8) \qquad \Delta(\hat{d}_i, \hat{d}_{i+1}) = \max_{d \in D'(\hat{d}_i)} \{\Delta(\hat{d}_i, d)\}$$

in $D'(\hat{d}_i)$. Since $D'(\hat{d}_{m(n-1)}) = \phi$ where $\hat{d}_{m(n-1)}(x_i) = a_n$ for all i, we have $N_1 \geqq m(n-1)+1$. Let us suppose that there existed another admissible rule $\hat{d}^*$ besides $\hat{d}_0, \cdots, \hat{d}_{m(n-1)}$. Then we can take out some $\hat{d}_i$ and $\hat{d}_{i+1}$ from among $\hat{d}_1, \cdots, \hat{d}_{m(n-1)}$ which satisfies

$$(9) \qquad R(\theta_1, \hat{d}_i) < R(\theta_1, \hat{d}^*) < R(\theta_1, \hat{d}_{i+1}) .$$

Since $\hat{d}_i$, $\hat{d}_{i+1}$ and $\hat{d}^*$ are admissible we have

$$(10) \qquad R(\theta_2, \hat{d}_i) > R(\theta_2, \hat{d}^*) > R(\theta_2, \hat{d}_{i+1}) .$$

Using (9) and (10), we get

$$\Delta(\hat{d}_i, \hat{d}^*) \geqq \Delta(\hat{d}_i, \hat{d}_{i+1}) .$$

This contradicts (8). Hence under the condition (7), $N_1 = m(n-1)+1$. If some quardruplet (i, i', k, k') satisfies (5), then there exist rules $\hat{d}_r$, $\hat{d}_{r+1}$ and $\hat{d}_{r+2}$, say, which satisfy

$$(11) \qquad (\hat{d}_r, \hat{d}_{r+1}) = (\hat{d}_{r+1}, \hat{d}_{r+2}) .$$

In fact, since

$$\begin{aligned}\Delta(\hat{d}_r, \hat{d}_{r+1}) &= \frac{R(\theta_2, \hat{d}_{r+1}) - R(\theta_2, \hat{d}_r)}{R(\theta_1, \hat{d}_r) - R(\theta_1, \hat{d}_{r+1})} \\ &= \frac{\{L(\theta_2, \hat{d}_{r+1}(x_k)) - L(\theta_2, \hat{d}_r(x_k))\} f(x_k | \theta_2)}{\{L(\theta_1, \hat{d}_r(x_k)) - L(\theta_1, \hat{d}_{r+1}(x_k))\} f(x_k | \theta_1)} \\ &= \nabla(i, 1, k)\end{aligned}$$

and

$$\Delta(\hat{d}_{r+1}, \hat{d}_{r+2}) = \nabla(i', 1, k'),$$

we get (11). It is easy to see that if $\hat{d}_r$, $\hat{d}_{r+1}$, $\hat{d}_{r+2}$ satisfy (11), then $\hat{d}_{r+1}$ can be expressed as a convex linear combination of $\hat{d}_r$ and $\hat{d}_{r+2}$. Hence $\hat{d}_{r+1}$ can not be an essentially nonrandomized rule. Therefore if t_1 quardruplets (i, i', k, k') satisfied (5), we get $N_1 = m(n-1)+1-t_1$.

(b) Consider the new problem (Decision problem B) with $L'(\theta_1, a_k) = -L(\theta_2, a_k)$, $L'(\theta_2, a_k) = -L(\theta_1, a_k)$, $f'(x_j|\theta_1) = f(x_j|\theta_2)$ and $f'(x_j|\theta_2) = f(x_j|\theta_1)$. By the definition of C_1 and C_2, a rule in C_1 of Problem A is a rule in C_2 of Problem B. Since in Problem B

$$\frac{L'(\theta_1, a_i)-L'(\theta_1, a_{i-1})}{L'(\theta_1, a_{i+1})-L'(\theta_1, a_i)} \cdot \frac{L'(\theta_2, a_{i+1})-L'(\theta_2, a_i)}{L'(\theta_2, a_i)-L'(\theta_2, a_{i-1})} > 1, \qquad \text{for } i=2, \cdots, n-1$$

by Theorem 4 in [2], all rules which call for a_i $(i=2, \cdots, n-1)$ are not admissible. Therefore Problem B reduces to a 2-action problem. Thus as in (a) we have $N_2 = m(2-1)+1-t_2 = m+1-t_2$ provided that t_2 doubles (k, k') satisfied (6).

(c) It is easy to see that $N_3 = N_1 + N_2 - 2 = mn - (t_1 + t_2)$.

Example. Consider the following problem with $L(\theta, a)$ and $f(x|\theta)$ given by table 1 and 2, respectively.

	a_1	a_2	a_3
θ_1	0	1	4
θ_2	5	3	2

Table 1: $L(\theta, a)$

	x_1	x_2	x_3	x_4
θ_1	0.40	0.30	0.20	0.10
θ_2	0.20	0.15	0.40	0.25

Table 2: $f(x|\theta)$

Since

$$\frac{L(\theta_1, a_2)-L(\theta_1, a_1)}{L(\theta_1, a_3)-L(\theta_1, a_2)} \cdot \frac{L(\theta_2, a_2)-L(\theta_2, a_3)}{L(\theta_2, a_1)-L(\theta_2, a_2)} = \frac{1}{6} < 1$$

and

$$\nabla(1, 2, 1)=1, \quad \nabla(1, 2, 2)=1, \quad \nabla(1, 2, 3)=4, \quad \nabla(1, 2, 4)=5,$$

$$\nabla(2, 3, 1)=\frac{1}{6}, \quad \nabla(2, 3, 2)=\frac{1}{6}, \quad \nabla(2, 3, 3)=\frac{2}{3}, \quad \nabla(2, 3, 4)=\frac{5}{6},$$

by Theorem 2, $N_1 = 4\times 2+1-2=7$, $N_2 = 4+1-1=4$ and $N_3 = 12-3=9$ and is also shown in Fig. 1. Furthermore by Theorem 1, we can get the following 11 nonrandomized rules in C_1 (See Fig. 2),

$$\hat{d}_0^* = (a_1, a_1, a_1, a_1)$$
$$\hat{d}_1^* = (a_1, a_1, a_1, a_2)$$
$$\hat{d}_2^* = (a_1, a_1, a_2, a_2)$$
$$\hat{d}_3 = (a_2, a_1, a_2, a_2)$$
$$\hat{d}_4 = (a_1, a_2, a_2, a_2)$$
$$\hat{d}_5^* = (a_2, a_2, a_2, a_2)$$
$$\hat{d}_6^* = (a_2, a_2, a_2, a_3)$$
$$\hat{d}_7^* = (a_2, a_2, a_3, a_3)$$
$$\hat{d}_8 = (a_3, a_2, a_3, a_3)$$
$$\hat{d}_9 = (a_2, a_3, a_3, a_3)$$
$$\hat{d}_{10}^* = (a_3, a_3, a_3, a_3)$$

where $*$ denote the rule is essentially nonrandomized rule.

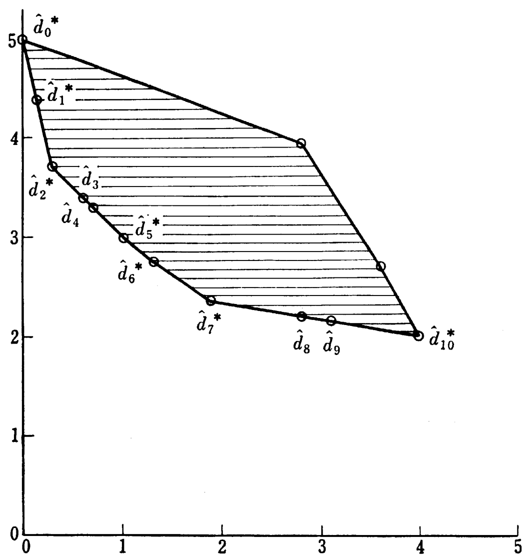

Fig. 1

Acknowledgement

The author wishes to thank the referees for their helpful comments in revising the paper.

THE INSTITUTE OF STATISTICAL MATHEMATICS

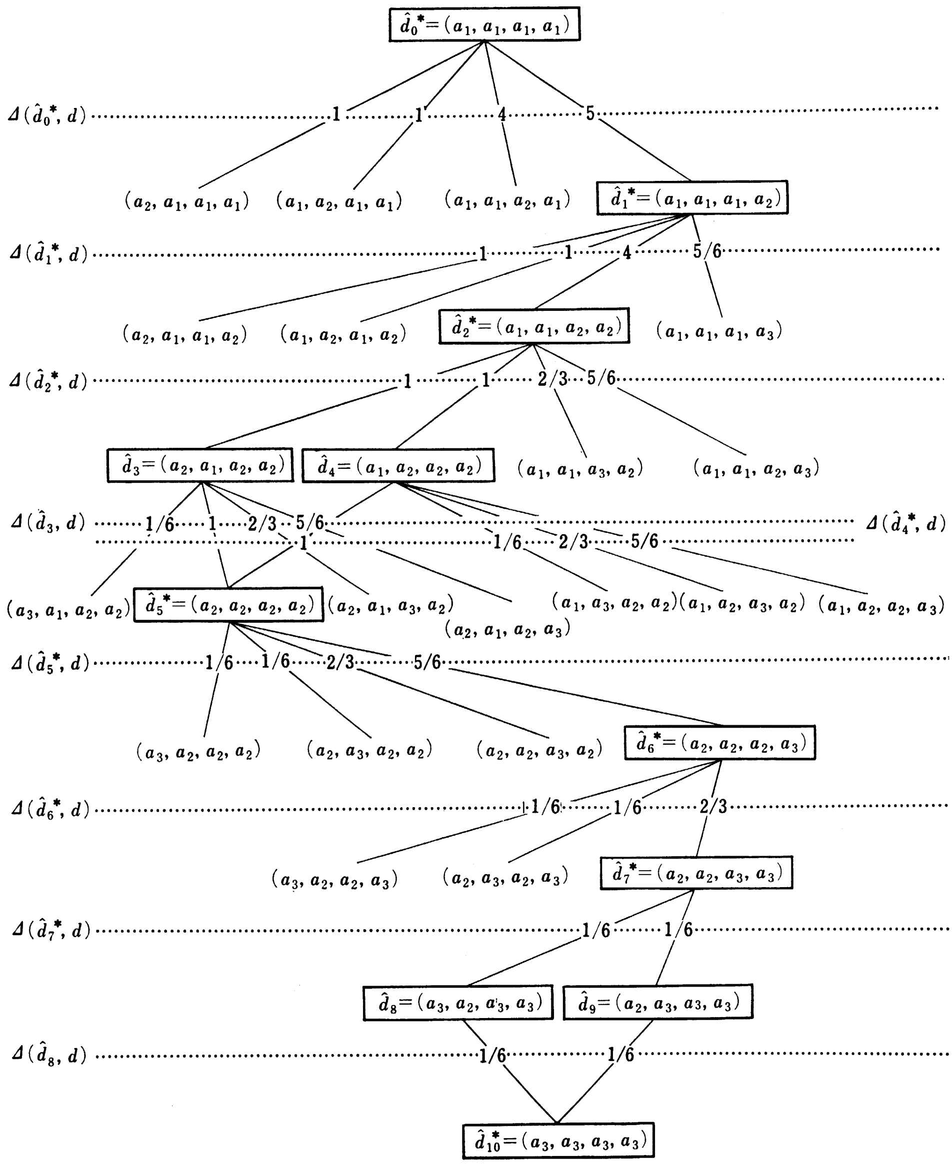

Fig. 2

REFERENCES

[1] Ferguson, T. S. (1967). *Mathematical Statistics: A Decision Theoretic Approach*, Academic Press, New York.

[2] Murakami, M. (1976). On the reduction to a complete class in multiple decision problems, *Ann. Inst. Statist. Math.*, **28**, A, 145-165.

Ann. Inst. Statist. Math.
36 (1984), Part A, 521

CORRECTIONS TO "SOME PROPERTIES OF THE RISK SET IN MULTIPLE DECISION PROBLEMS"

MASAKATSU MURAKAMI

In the above titled paper (this Annals Vol. 35, No. 2, A, (1983), pp. 175–183), the following corrections should be made:

On page 179, line 15 from the bottom

$$\nabla(i, j, k)=\frac{\{L(\theta_2, a_i)-L(\theta_2, a_j)\}f(x_k|\theta_2)}{\{L(\theta_1, a_j)-L(\theta_1, a_i)\}f(x_k|\theta_1)}$$

$$\Longrightarrow \nabla(i, j, k)=\frac{\{L(\theta_2, a_i)-L(\theta_2, a_{i+j})\}f(x_k|\theta_2)}{\{L(\theta_1, a_{i+j})-L(\theta_1, a_i)\}f(x_k|\theta_1)} .$$

On page 179, in (6)

"n" should be "$n-1$".

THE INSTITUTE OF STATISTICAL MATHEMATICS