A Sequential Multiple Pattern Recognition Plan Based on Two-Sample Rank Test

Geng, Zhi
Department of Information Systems, Interdisciplinary Graduate School of Engineering Sciences, Kyushu University

Asano, Choichiro
Department of Information Systems, Interdisciplinary Graduate School of Engineering Sciences, Kyushu University

https://doi.org/10.15017/17647

出版情報：九州大学大学院総合理工学報告. 7 (2), pp.285-297, 1986-01-01. 九州大学大学院総合理工学研究科

A Sequential Multiple Pattern Recognition Plan Based on Two-Sample Rank Test

Zhi Geng* and Chooichiro ASANO**

(1985, 9, 30)

Several sequential multiple pattern recognition plans are proposed in this paper in view of rank test for Lehmann alternatives. There exist two ways of ranking measurements (i) ranking measurements within-group, and (ii) reranking measurements at each stage of the sequential process. Moreover, there exist two sequential types (a) sequential observation vectors with p features, and (b) sequential measurements for p features in a cyclic order. Four plans are formulated on the basis of the possible combinations of (i), (ii) and (a), (b). Numerical properties of the plans are concretely illustrated with various values of designing constants by simulations.

1. Introduction

The plans of pattern recognition system in a random environment have been developed from the practical and theoretical point of view. In this regard, there exist some preferential requirements as follows,

(i) the plan is of a sequential type,
(ii) the recognition is based on multiple features, if necessary,
(iii) the assumption is minimal on the distribution law of features,
(iv) the plan is of a powerful discrimination, and
(v) the rule of the procedure is rather simple.

Let the present sequential plan be a comparative sequential group sampling for two populations, a test population T and the control C, with p features, i.e. so-called a statistically sequential pattern matching system. According to the sequential scheme with p features, $p \geq 2$, there are two types, regarding what is sequential. The one is closely related to p features, and p feature measurements are sequentially observed like $X_1 \rightarrow X_2 \rightarrow \cdots \rightarrow X_p \rightarrow X_1 \rightarrow X_2 \rightarrow \cdots$ cyclically, and each size of observations at a stage is specified effectively in view of cost. The other type is related to three sequential sampling plans, i.e. one-by-one, group and partial samplings. These are essentially considered on designing an optimal sequential plan under both costs of sampling and sample.

In the universal situations without any prior information on the distribution laws of feature measurements, nonparametric tests are appropriate. Especially the most powerful rank test for Lehmann alternatives may be preferable to another nonparametric tests. Naturally, there exist real questions to Lehmann alternatives, but some positive merits are also theoretically pointed out, e.g. Wilcoxon-Rhodes-Bradley [1], Savage [2]. Also, practically Fu [3] has written as "Lehmann alternatives have been found to provide the typical de-

* Department of Information Systems, Graduate Student
** Department of Information Systems

viations which usually prevail in many probability distributions describing various pattern classes".

In the present paper, from the viewpoint of the whole requirements (i)-(iv), several sequential multiple pattern recognition plans are proposed comparatively for testing whether both populations C and T are equal or not , and may be called a set of sequential group sampling multi-feature two-sample most powerful rank tests to Lehmann alternatives. Since the present paper gives precisely their statistical properties, a certain optimal plan may be given in practice.

2. The pattern recognition plans

Let us give a sequential pattern recognition plan with multi-feature rank test and then several modifications of the plan. Suppose that observation vector $\mathbf{X}^{(i)}_{jt} = [x^{(i)}_{jt1}, x^{(i)}_{jt2}, \cdots, x^{(i)}_{jtn_t}]'$ is obtained on j-th feature variable at t-th stage from i-th population, where the elements are mutually independent and $i = 1, 2$. Then the pattern matching is now to test a null hypothesis that two distributions are similar, against the Lehmann alternatives.

Simply writing $\mathbf{X}_{jt}$ for $\mathbf{X}^{(1)}_{jt}$ from C, and $\mathbf{Y}_{jt}$ for $\mathbf{X}^{(2)}_{jt}$ from T, let $F_j\ (x_{jtk})$ and $G_j\ (y_{jtk})$ be the cumulative distribution functions $\Pr\ (X_j \leqq x_{jtk})$ of X_j and $\Pr\ (Y_j \leqq y_{jtk})$ of Y_j, respectively. Then a null hypothesis H_0 and the alternatives H_1 are

$$H_0 : G_j\ (Y) = F_j\ (X), \quad H_1 : G_j\ (Y) = F_j^{r_j}\ (X), \quad r_j > 1, \quad j = 1, 2, \cdots, p.$$

This implies that $\Pr\ (X_j \leqq Y_j) = r_j/(r_j + 1)$. The test is also equivalent for $H_0 : r_j = 1$ and $H_1 : r_j > 1$, $j = 1, 2, \cdots, p$. Then the sequential test is given with Wald's [4] upper and lower horizontal boundary values A and B, where $A = (1-\beta)/\alpha$ and $B = \beta/(1-\alpha)$.

According to the above procedure , if no decision, i.e. neither H_0 nor H_1, is obtained, successively two observation vectors $\mathbf{X}^{(i)}_{21}$ with n_2 elements, $i = 1, 2$, are observed on the second feature X_2, and the similar testing hypothesis is continued. Thus so far as the decision is reserved, p feature measuremens are subsequently observed in a cyclic order. The numbers of mesurements are fixed for the respective features, say n_i for X_i, $i = 1, 2, \cdots, p$.

Concerning the comparative situation of pattern recognition, there exist two cases on the test procedure, i.e.

(i) never combining $\mathbf{X}^{(i)}_{j1}, \mathbf{X}^{(i)}_{j2}, \mathbf{X}^{(i)}_{j3}, \cdots$

(ii) always combining $\mathbf{X}^{(i)}_{j1}, \mathbf{X}^{(i)}_{j2}, \mathbf{X}^{(i)}_{j3}, \cdots$

for $i = 1, 2$ and $j = 1, 2, \cdots, p$.

If such observation vectors are suffered environmental block-effects like in simultaneous biological assay, the above case (i) is valid, and if observation vectors are obtained in a well-controlled environment, the case (ii) is preferable in the sense of statistical analysis.

Moreover, there exist two sequential types, consisting of

(a) sequential observation vectors with p features, and

(b) sequential measurements for p features in a cyclic order,

where either (a) or (b) depends on the physical conditions in practice.

3. Numerical illustration of the procedure

To clarify the above procedure, let us illustrate a simple pattern recognition plan with a group sampling rule for a feature.

The procedure is to test a hypothesis H_0: $G=F$ against H_1: $G=F^r$ with $r=3$ and $\alpha=\beta=0.1$. The size m of group sampling at every stage of sequential sampling is now three. Thus both boundaries, upper A and lower B, are given as $A=9.0$ and $B=0.11$. Let the successive measurements be x_1, x_2, ⋯ and y_1, y_2, ⋯ for both populations C and T, independently and respectively, as shown in **Table 1**.

Table 1. Source data for illustrations

(X):	0.31	0.25	0.47		0.35	0.30	0.21		0.38	0.43	0.22		0.34	0.37	0.24	
				→				→				→				→⋯
(Y):	0.49	0.36	0.42		0.39	0.45	0.28		0.41	0.48	0.40		0.44	0.46	0.33	

Table 2. An illustrative procedure at 1-st stage

k=6	$\{V_i\}$:	0.25	0.31	0.36	0.42	0.47	0.49
rank	$\{R_i\}$:	1	2	3	4	5	6
score	$\{S_i\}$:	1	1	r	r	1	r
	$B<\lambda_6=2.250<A$						

Table 3 An illustration for type (i) at 2-nd stage

k=12	$\{V_i\}$:	0.25	0.31	0.36	0.42	0.47	0.49	0.21	0.28	0.30	0.35	0.39	0.45
rank	$\{R_i\}$:	1	2	3	4	5	6	1	2	3	4	5	6
score	$\{S_i\}$:	1	1	r	r	1	r	1	r	1	1	r	r
	$B<\lambda_{12}=3.375<A$												

Table 4. An illustration for type (i) at 4-th stage

k=24	$\{V_i\}$:	0.25	0.31	0.36	0.42	0.47	0.49	0.21	0.28	0.30	0.35	0.39	0.45
		0.22	0.38	0.40	0.41	0.43	0.48	0.24	0.33	0.34	0.37	0.44	0.46
rank	$\{R_i\}$:	1	2	3	4	5	6	1	2	3	4	5	6
		1	2	3	4	5	6	1	2	3	4	5	6
score	$\{S_i\}$:	1	1	r	r	1	r	1	r	1	1	r	r
		1	1	r	r	1	r	1	r	1	1	r	r
	$\lambda_{24}=11.39>A$ (to accept H_1).												

According to the sequential group sampling rule with the size three, we obtain three samples from each population at the first stage and observe the measurements. **Table 2** shows the ranks and scores of six observations at this stage, and the value of the likelihood ratio is 2.250, which locates between A and B. Therefore, any decision has to be reserved. Thus we have to continue the sequential sampling and take additional feature measurements.

At the second stage, we observe three samples from each population, with either situa-

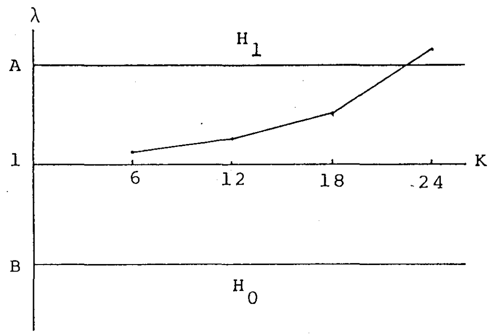


Figure 1. An illustrative path for un-combining

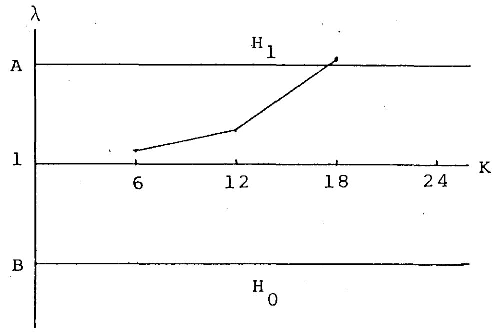


Figure 2. An illustrative path for combining measurements

Table 5. An illustration for type (ii) at 3-rd stage

k=18	$\{V_i\}$:	0.21	0.22	0.25	0.28	0.30	0.31	0.35	0.36	0.38	0.39	0.40	0.41
		0.42	0.43	0.45	0.47	0.48	0.49						
rank	$\{Ri\}$:	1	2	3	4	5	6	7	8	9	10	11	12
		13	14	15	16	17	18						
score	$\{Si\}$:	1	1	1	r	1	1	1	r	1	r	r	r
		r	1	r	1	r	r						
		$\lambda_{18} = 9.909 > A$ (to accept H_1).											

tion described at (i) or (ii) in section 2. For (i), only the second six samples from both populations are combined, ranked and scored for the respective groups. The value of the likelihood ratio is 3.375 as shown in **Table 3**. Thus the next stage is needed. Such stages are continued until a decision, H_0 or H_1, is given. At the fourth stage, since the value of likelihood ratio is larger than A, a decision H_1 is accepted as shown in **Table 4**.

On the other hand, for (ii) new observations are always combined to all of previous observations, ranked and scored for the group measurements. At the second stage, we obtain a value 3.938 as the likelihood ratio, which indicates the necessity of more stages. **Table 5** shows the result obtaining a decision H_1. **Figures 1** and **2** show the summarized paths for (i) and (ii), respectively.

4. Mathematical Formulation of the Plans

In this section, four plans are formulated on the basis of the possible combination of (i), (ii) and (a), (b) in section 2. Namely the respective likelihood ratio functions, the operating characteristic functions and the average sample numbers are given, under the condition that all features of a pattern are independent to test a hypothesis H_0: $G_j = F_j$ against the alternative H_1: $G_j = F_j^{r_j}$, $r_j > 1$, $j = 1, 2, \cdots, p$.

4.1 The plan with (i) and (a)

The first plan is not to combine measurements, i.e. only to rank measuements obtained at each stage, and is to observe simultaneously all features at each stage.

4.1.1 Likelihood ratio function and decision

Let p measurements be $(x_{jk1}, \cdots, x_{jkn_j}, y_{jk1}, \cdots, y_{jkn_j})$, $j=1, 2, \cdots, p$, $k=1, 2, \cdots, q$, independently observed from two populations C and T at k-th stage, and belong to $F_j(X)$ and $G_j(Y)$, respectively. These measurements obtained at each stage till q-th stage are combined and ordered increasingly. Then the joint probability to show the following p sequential rank vectors is

$$P(R)=P(v_{jk1}<\cdots<v_{jk(2n_j)}\ ,\ j=1, \cdots, p\ ,\ k=1, \cdots, q)$$

$$=\int_R\cdots\int \prod_{k=1}^{q}\prod_{j=1}^{p}\prod_{i=1}^{2n_j} dy_{jki}^{s_{jki}}=\prod_{k=1}^{q}\prod_{j=1}^{p}\prod_{i=1}^{2n_j}(s_{jki}/\sum_{m=1}^{i}s_{jkm}),$$

where
$$y_{jki}^{s_{jki}}=\begin{cases}F_j(x), & v_{jkm}\in C,\\ G_j(x)=F^{r_j}(x), & v_{jkm}\in T,\end{cases}\qquad \text{and}\quad s_{jkm}=\begin{cases}1, & v_{jkm}\in C,\\ r_j, & v_{jkm}\in T..\end{cases}$$

The Likelihood ratio is given as

$$\lambda_q=P(R|H_1)/P(R|H_0)=\prod_{k=1}^{q}\prod_{i=1}^{p}r_i^{n_i}(2n_i)!\prod_{j=1}^{2n_i}(1/\sum_{m=1}^{j}s_{ikm})$$

and the boundaries are $A=(1-\beta)/\alpha$ and $B=\beta/(1-\alpha)$.

Then the decision rule is to accept H_0, if $\lambda_q\leqq B$, and to accept H_1, if $\lambda_q\geqq A$, and to continue sampling at the next stage, if $B<\lambda_q<A$.

4.1.2 OC function

The operating characteristic function is approximately given by $L\fallingdotseq(A^h-1)/(A^h-B^h)$, under $E(e^{zh})=1$, $h\neq 0$, where h is essentially a function of an alternative r_i and R_i for the operating characteristic curved surface, $i=1, 2, \cdots, p$ through a random variable Z.

To evaluate h actually, letting R be a set of all possible arrangements of ranks $\{v_{it}, t=1, \cdots, 2n_i, i=1, \cdots, p\}$ and similarly R′ be only for C, the following relation is to be applied to compute iteratively a convergent value of h.

$$E(e^{zh})=\sum_R\left\{\prod_{i=1}^{p}\prod_{j=1}^{2n_i}(jr_i^{1/2}/\sum_{t=1}^{j}s_{it})\right\}^h P(R)$$

$$=\sum_{R'}\prod_{i=1}^{p}(n_i!)^2\prod_{j=1}^{2n_i}\left\{(jr_i^{1/2}/\sum_{t=1}^{j}s_{it})^h/(s'_{ij}/\sum_{t=1}^{j}s'_{it})\right\}=1,$$

where
$$s'_{it}=\begin{cases}1 & v_{it}\in C,\\ R_i & v_{it}\in T,\end{cases}\qquad s_{it}=\begin{cases}1 & v_{it}\in C,\\ r_i & v_{it}\in T.\end{cases}$$

4.1.3 ASN function

The ASN function is approximately expressed by Wald's principle in the following way.

$$E(q)\fallingdotseq\{L\log B+(1-L)\log A\}/E(Z),$$

$$\text{where}\quad E(Z) = \sum_{R'} \left\{ \log \left(\prod_{i=1}^{p} (2n_i)!\, r_i^{n_i} / \prod_{j=1}^{2n_i} \sum_{m=1}^{j} s_{im} \right) \right\} P(R')$$

$$= \sum_{R} \left\{ \log \left(\prod_{i=1}^{p} (2n_i)!\, r_i^{n_i} / \prod_{j=1}^{2n_i} \sum_{m=1}^{j} s_{im} \right) \right\} \prod_{i=1}^{p} \left\{ (n_i!)^2 R_i^{n_i} / \prod_{j=1}^{2n_i} \sum_{m=1}^{j} s'_{im} \right\}$$

4.2 The plan with (i) and (b)

This plan is also not to combine measurements, but is sequentially and cyclically to observe p features.

4.2.1 Likelihood ratio function

Let $(x_{jm1}, \ldots, x_{jmn_j}, y_{jm1}, \ldots y_{jmn_j})$, $j=1, \cdots, p$, $m=1, \cdots, q-1$, and $(x_{rq1}, \ldots, x_{rqn_r}, y_{rq1}, \ldots, y_{rqn_r})$, $r=1, 2, \cdots, k$ $(<p)$, be sequential measurements from two populations till k-th feature at q-th stage, and let them be ordered increasingly as follows: R: $v_{jm1} < \cdots < v_{jm(2n_j)}$, $m=1, 2, \cdots q$, $m<q : j=1, 2, \cdots, p$, and $m=q : j=1, 2, \cdots, k$.

Since the probability for R is

$$P(R) = \prod_{i=1}^{p} \prod_{m=1}^{q-\delta_i} \prod_{j=1}^{2n_i} (j r_i^{1/2} / \sum_{t=1}^{j} s_{imt}), \quad \text{where } \delta_i = \begin{cases} 0, & i \leqq k \\ 1, & i > k, \end{cases}$$

the present likelihood ratio is given by

$$\lambda = \frac{P(R|H_1)}{P(R|H_0)} = \prod_{i=1}^{p} \prod_{j=1}^{q-\delta_i} \left\{ (2n_i)!\, r_i^{n_i} / \prod_{t=1}^{2n_i} \sum_{m=1}^{t} s_{ijm} \right\}.$$

4.2.2 OC function

The operating characteristic function is given as follows,

$$L = \sum_{v=1}^{\infty} P(\lambda_v \leqq B), \qquad \text{where } v = p(q-1)+k, \quad \delta_v = \begin{cases} 0, & k \leqq p \\ 1, & k > p \end{cases}$$

$$= \sum_{v=1}^{\infty} P\left\{ \frac{1}{B_v} \sum_{i=1}^{p} \sum_{j=1}^{q-\delta_v} (d_{ij} - \mu_i) \geqq \frac{1}{B_v} (C_v - (q-\delta_v) \sum_{i=1}^{p} \mu_i) \right\},$$

$$\text{where}\quad \mu_i = E(d_{ij}) = \sum_{R} (n_i!)^2 \prod_{t=1}^{2n_i} (s'_{ijt} / \sum_{m=1}^{t} s'_{ijm})\, d_{ij},$$

$$d_{ij} = \sum_{t=1}^{2n_i} \log \left(\sum_{m=1}^{t} s_{ijm} / t \right), \quad s'_{ijm} = \begin{cases} 1, & v_{ijm} \in C, \\ R_i, & v_{ijm} \in T, \end{cases}$$

$$B_v^2 = \sum_{i=1}^{p}(q-\delta_v)\left\{\sum_R (d_{ij}-\mu_i)^2 (n_i!)^2 \prod_{t=1}^{2n_i}(s'_{ijt}/\sum_{m=1}^{t} s'_{ijm})\right\},$$

$$C_v^2 = -\log B + \sum_{i=1}^{p}\sum_{j=1}^{q-\delta_v}\sum_{t=1}^{2n_i}\log r_i^{1/2},$$

and where R is a set of all possible arrangements of ranks, r_i's are the alternatives, and R_i's are arbitrary parameters for the operating characteristic curved surface.

Furthermore, according to Liapounoff theory, the above function has a property such that

$$\lim_{v\to\infty} P\left\{\sum_{k=1}^{v}(d_{kj}-\mu_k)/B_v \leq a\right\} = \int_{-\infty}^{a}\exp(-Z^2/2)\,dz/\sqrt{2\pi}\ , \quad \text{where } B_v^2=\sum_{k=1}^{v}\sigma_k^2,$$

since Lindeberg-Feller condition is satisfied as follows,

$$\frac{1}{B_v^{2+\delta}}\sum_{i=1}^{p}(q-\delta_v)\left\{\sum_R (d_{ij}-\mu_i)^{2+\delta}P(R)\right\} \longrightarrow 0, \quad \text{for } \exists\,\delta>0,\ v\longrightarrow\infty.$$

4.2.3 ASN function

This plan is to observe sequentially and cyclically the measurements like $x_1 \to x_2 \to \cdots \to x_p \to x_1 \to x_2 \to \cdots$, if no decision is obtained. For k-th feature at m-th stage, we can obtain generally the ASN function as follows,

$$\sum_{i=1}^{E(k)} E(m)E(z_i) + \sum_{i=E(k)+1}^{p} E(m-1)\,E(z_i) = L\log B + (1-L)\log A$$

$$\text{where } z_i = \sum_{j=1}^{2n_i}\left\{\log\,(jr_i^{1/2}/\sum_{t=1}^{j} s_{it})\right\}, \quad \text{and } E(z_i) = \sum_R z_i\,(n_i!)^2 \prod_{j=1}^{2n_i}(s'_{ij}/\sum_{t=1}^{j} s'_{it}),$$

$$\text{for} \quad s_{ij} = \begin{cases} 1, & v_{ij}\in C, \\ r_i, & v_{ij}\in T, \end{cases} \qquad s'_{ij} = \begin{cases} 1, & v_{ij}\in C, \\ R_i, & v_{ij}\in T. \end{cases}$$

4.3 The plan with (ii) and (a)

This plan is to combine measurements for each feature till then, reranking all measurements, and sampling is sequential to observe simultaneously all p features.

4.3.1 Likelihood ratio function

Let $x_{jm1}, \dots, x_{jmn_j}, y_{jm1}, \dots, y_{jmn_j}$ be measurements till q-th stage, and be ordered increasingly for each feature as follows, R: $v_{j*1} < \dots < v_{j*2qn_j}$, $j=1, 2, \cdots, p$, $m=1, 2, \cdots, q$, where $*$ means that the measurements of the j-th feature are reranked at every stage.

Then the probability for the arrangement of R under the assumption C=T is shown as follows,

$$P(R) = \prod_{i=1}^{p} (r_i^{mn_i} / \prod_{k=1}^{2mn_i} \sum_{t=1}^{k} s_{it}), \quad \text{where } s_{it} = \begin{cases} 1, & v_{i*t} \varepsilon C \\ r_i, & v_{i*t} \varepsilon T. \end{cases}$$

Therefore, the likelihood ratio function can be shown by

$$\lambda = \frac{P(R|H_1)}{P(R|H_0)} = \prod_{i=1}^{p} \left\{ (2mn_i)!\, r_i^{mn_i} / \prod_{k=1}^{2mn_i} \sum_{t=1}^{k} s_{it} \right\}.$$

4.3.2 OC function

The OC function can be similarly given as follows,

$$L = \sum_{R^*} P(R_i, i=1,\cdots, p) = \sum_{R^*} \prod_{i=1}^{p} \prod_{k=1}^{2mn_i} (s'_{ik} / \sum_{t=1}^{k} s'_{it})$$

$$\text{where} \quad s'_{it} = \begin{cases} 1, & v_{i*t} \varepsilon C, \\ R_i, & v_{i*t} \varepsilon T, \end{cases} \qquad R^* = \left\{ R |\, \lambda = P(R|H_1)/P(R|H_0) \leqq B \right\}.$$

4.3.3 ASN function

The ASN function is also approximately given by

$$ASN \fallingdotseq \sum_{R^*} L_{r*}\, P(R_i, i=1, 2, \cdots, p)$$

where $R^* = \left\{ R |\, \lambda = P(R|H_1)/P(R|H_0) \leqq B, \text{ or } \geqq A \right\}$, and L_{r*} means the number of elements in $r^* \varepsilon R^*$.

4.4 The plan with (ii) and (b)

Since this plan is similar in section 4.3, excepting the sequential measurements on p features, the properties of the plan are briefly shown in the following way.

4.4.1 Likelihood ratio function

$$\lambda = \prod_{i=1}^{p} \left\{ (2m_i)!\, r_i^{m_i} / \prod_{k=1}^{2m_i} \sum_{m=1}^{k} s_{im} \right\},$$ where $2m_i$ is the total number of $\{x_i\}$ and $\{y_i\}$.

4.4.2 OC function

$$L = \sum_{R} \prod_{i=1}^{p} (\prod_{k=1}^{2m_i} s'_{ik} / \prod_{k=1}^{2m_i} \sum_{t=1}^{k} s'_{it}).$$

4.4.3 ASN function

$$ASN = \sum_{R^*} L_{r*} P(R\, R \varepsilon R^*).$$

4.5 Generalized Likelihood Ratio Function for the Plans

Let us show a generalization for four plans in the previous sections. Let an example be shown at first to introduce the idea.

(i) An example

Suppose that ten feature variables $\{X_i\}$ are independently belonging to the respective distribution law $\{F_i\}$ and alternatively to $\{G_i\}$, where G_i is assumed to be $G_i = F_i^{r_i}$, $r_i > 1$, for $i = 1, 2, \cdots, 10$. Let $\{X_1, X_2, X_3, X_4, X_5\}$, $\{X_6, X_7, X_8\}$ and $\{X_9, X_{10}\}$ be assumed to be comparable. Then the probability of an event $A = (X_3 < X_2, X_4 < X_2, X_2 < X_1, X_5 < X_1, X_8 < X_6, X_7 < X_6, X_{10} < X_9)$ under H_1 is shown by

$$\Pr(A) = \left\{ \prod_{i=1}^{5} r_i / [r_3 r_4 r_5 (r_2 + r_3 + r_4)(r_1 + r_2 + r_3 + r_4 + r_5)] \right\}$$

$$\cdot \left\{ \prod_{i=6}^{8} r_i / [r_7 r_8 (r_6 + r_7 + r_8)] \right\} \cdot \left\{ r_9 r_{10} / [r_{10} (r_9 + r_{10})] \right\}.$$

This probability can be expressed by a tree expression in **Figure 3.** Basing on the tree expression, we express the probability simply by $P(A) = r_1 r_2 \cdots r_{10} / \alpha_1 \alpha_2 \cdots \alpha_{10}$, where α_i denotes a sum of the lower r_i's, i.e. $\alpha_1 = r_1 + r_2 + \cdots + r_5$, $\alpha_2 = r_2 + r_3 + r_4$, $\alpha_3 = r_3$, $\cdots$.

(ii) Generalized Formula

Let us show the generalized likelihood ratio test of a hypothesis H_0 against H_1, where

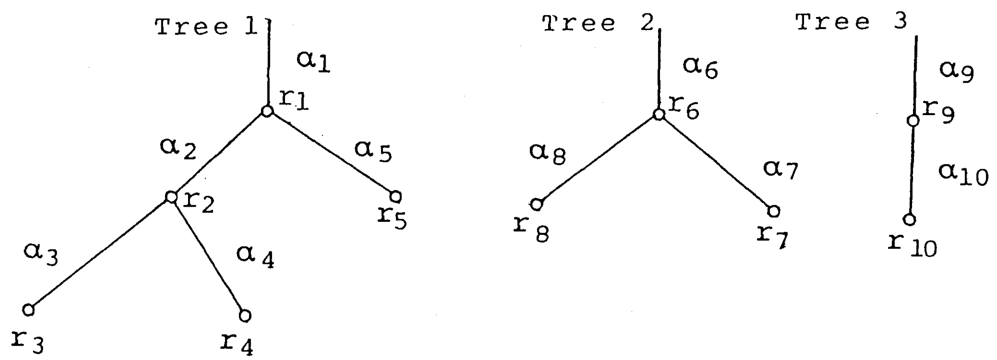


Figure 3. An illustrative tree-expression

$$H_0\colon G_i = F_{ir}^{\vartheta_i},\ H_1\colon G_i = F_i^{r_i},\ r_i > \vartheta_i,\ i = 1, 2, \cdots, p.$$

The generalized comparison tree is now shown as **Figure 4**. The probability of comparison tree for H_0 is given as

$$P_{H_0}(\text{Trees}) = \prod_{k=1}^{n} \prod_{i \in \text{Tree}_k} (\vartheta_i / \alpha_i), \quad \text{for } \alpha_i = \Sigma\, \vartheta_j.$$

where α_i is the sum of ϑ_j 's under the branch α_i in **Figure 4**. On the other hand, the probability of comparison tree for H_1 is given as

$$P_{H_1}(\text{Trees}) = \prod_{k=1}^{n} \prod_{i \in \text{Tree}_k} (r_i / \alpha_i),$$

where α_i is the sum of r_j 's under the branch α_i. Thus the likelihood ratio function is given by $\lambda = P_{H_1}/P_{H_0}$. Note that by replacing r_i with cr_i, and ϑ_i with $c\,\vartheta_i$, the likelihood ratio is invariant.

5. Designing the plans

In order to design four plans in the previous section, numerical evaluation of their properties is necessary, although the formulae are rather complicated, excepting some simple cases. For the sake of these realization, the following three methods are possible to apply

(i) computing the exact formula by a software program iteratively,
(ii) basing on computer simulations,
(iii) using asymptotic theories.

The iterative method of computing (i) may be appropriate to OC functions and ASN formulae given in section 4.3 and 4.4. The method of asymptotic theories (iii) may be appropriate to those in section 4.1 and 4.2. In the following section, numerical properties of the plans are concretely illustrated for various values of constants.

6. Evaluation of Sequential Plans

In order to evaluate performances of these plans with constants to be considered, sever-

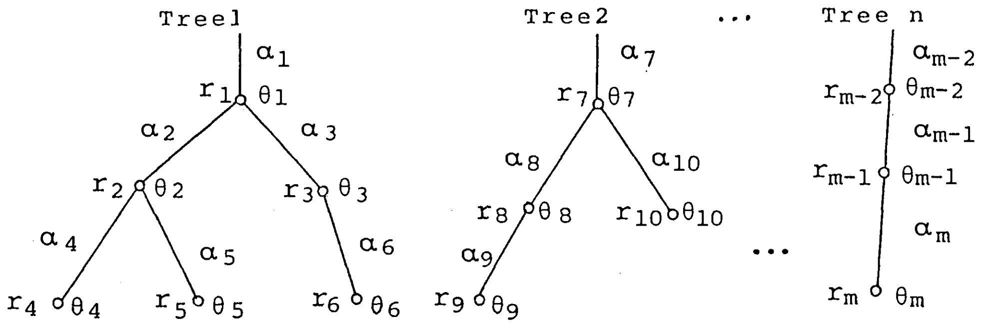
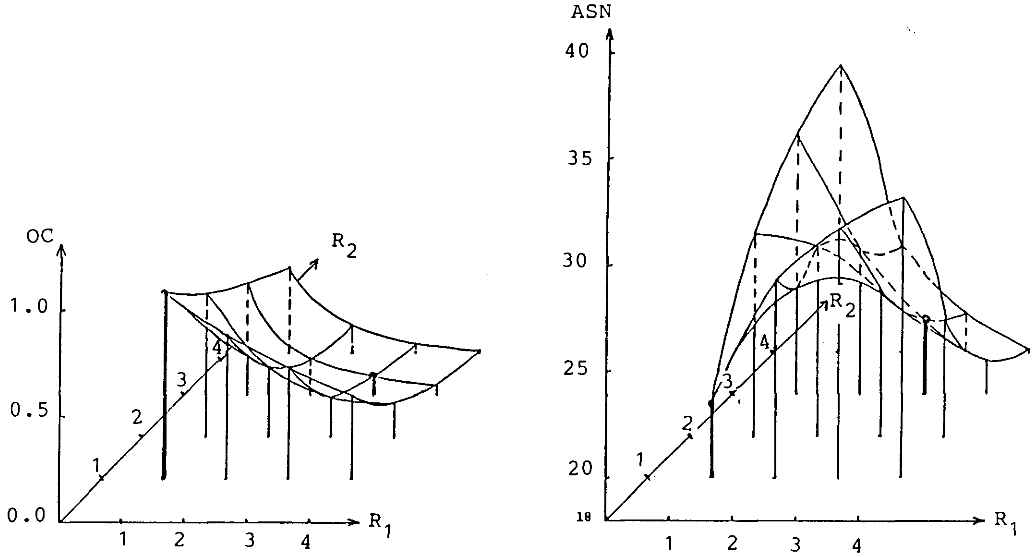


Figure 4. Generalization of tree expression

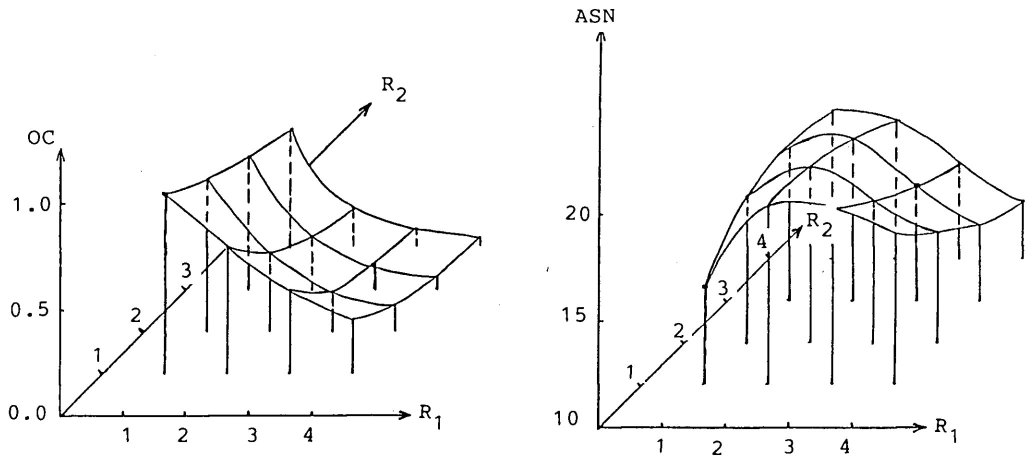
al plans were precisely simulated 5000 times at each lattice point on an M-240H (HITAC) computer.

In **Figures 5** and **6**, OC and ASN surfaces are shown for plans in sections 4.1 and 4.4. It may be clear in **Figure 5 (a)** and **Figure 6 (a)** that both OC surfaces are alike in their



(a) OC curve　　(b) ASN curve

Figure 5. An illustrative plan with (i) and (a)
$\alpha=\beta=0.2$, $n_i=3$, $r_i=3.0$, $i=1, 2$.



(a) OC curve　　(b) ASN curve

Figure 6. An illustrative plan with (ii) and (b)
$\alpha=\beta=0.2$, $n_i=3$, $r_i=3.0$, $i=1, 2$.

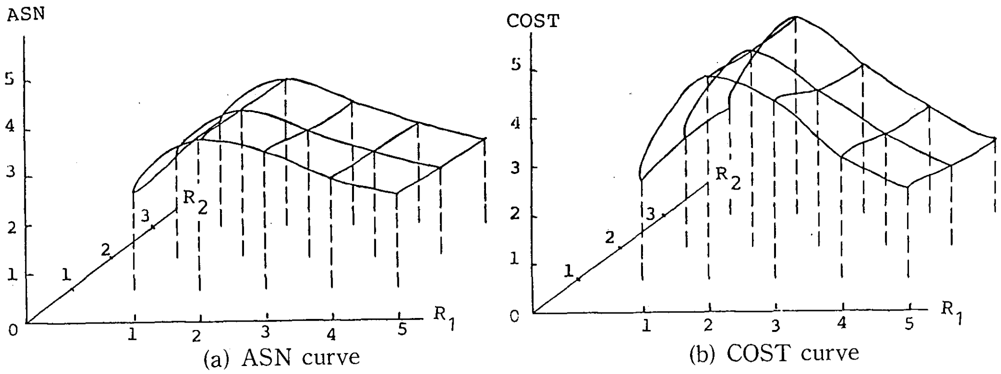


Figure 7. Properties of an illustrative plan with (ii) and (b), $p=2$, $r_1=4$, $n_1=5$, $c_1=5$ and $c_2=5$ for first feature, $r_2=2$, $n_2=2$, $c_1=4$ and $c_2=2$ for the second feature.

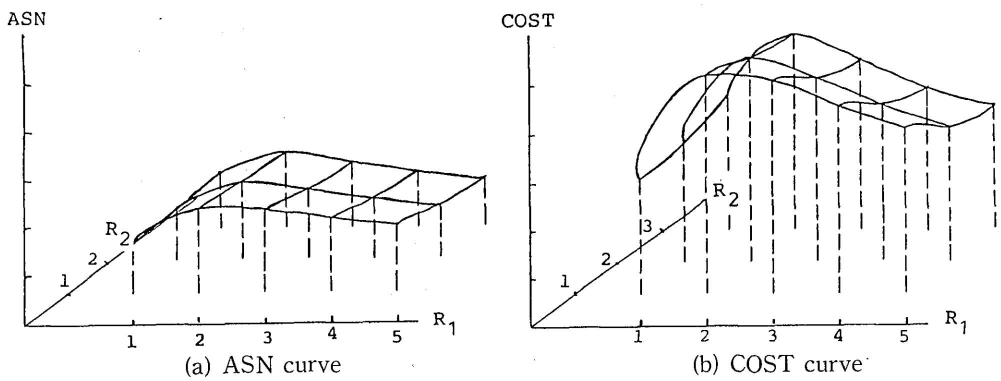


Figure 8. Properties of an illustrative plan with (ii) and (b), $p=2$, $r_1=4$, $n_1=2$, $c_1=5$, $c_2=1$ for the first feature, $r_2=2$, $n_2=1$, $c_1=4$ and $c_2=2$ for the second feature.

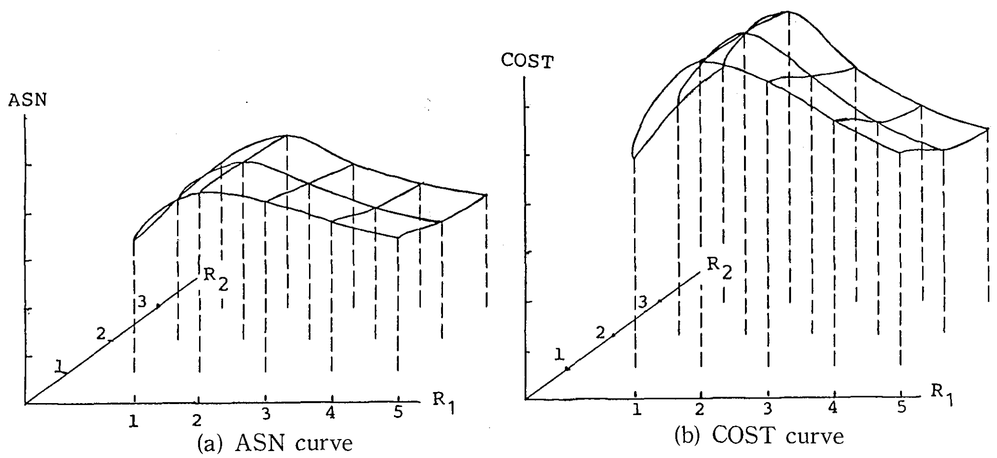


Figure 9. Properties of an illustrative plan with (ii) and (b), $p=2$, $r_1=2$, $n_1=2$, $c_1=4$, $c_2=2$ for the first feature, $r_2=5$, $n_2=4$, $c_1=5$ and $c_2=1$ for the second feature.

locations and shapes, although the ASN surface of plan in section 4.4 is lower than in section 4.1. This is because plans without combining measurements may be less infomation than plans with combining measurements, and obviously plans of sequential sampling with p simultaneous features come to more measurements than those of sequential feature sampling.

In order to design an optimal sequential group sampling plan with various types proposed in the present paper, cost performances of plans are investigated with their statistical properties. Namely, an illustrative interest is in how to determine the size of group sampling and the order of features on observation. In **Figures 7**, **8** and **9**, it becomes clear by comparing **Figure 7** with **Figure 8** that the ASN function with smaller group size is lower, but the cost function is larger, and by comparing **Figure 7** with **Figure 9**, that different order of features comes to different values of ASN and COST surfaces. Therefore, in view of cost performance under the general condition with sampling cost C_1 and sample cost C_2, there exists an optimal group size and an optimal order of features observed. In practice, designing an optimal sequential plan is very impotant. It is our recommendation at present that a computer simulation is easy to realize an optimal plan.

References

1) Wilcoxon, F., Rhodes, L. J. and Bradley, R. A. (1963). Two sequential two-sample grouped rank tests with applications to screening experiments, Biometrics, vol. 19, pp. 58−84.
2) Savage, I. R. (1980). Lehmann alternatives, Nonparametric statistical Inference, North-Holland.
3) Fu, K. S. (1963). Sequential Methods in Pattern Recognition and Machine Learning, Academic press.
4) Wald, A. (1947). Sequential Analysis, John-Wiley.