THE EXTREME VALUE IN A MULTIVARIATE NORMAL SAMPLE.

By

Akio KUDÔ

(Received July 20, 1957)

§ 1. Introduction

The theory of extreme value has some special importance in the theory of order statistics, and the testing of outlying observation may be considered to be the testing of the extreme value. (C. f. GUMBEL [1] p. 3.) The role of the test is to find out a discordant observation from a sample and reject it so that the inference based on this sample may be more accurate. Therefore, the author would like to mention incidentally, this test may be more adequately treated as one of the preliminary tests in the theory of successive process of statistical inferences in the terminology of T. KITAGAWA [2].

The same problem should arise in the case of multivariate sample and here what we should at first decide is what is the extreme value in a multivariate sample and secondly we have to test if the extreme value may be considered to be discordant.

In this paper we propose a test which is based on the statistic defined by $\text{Max}_\nu \sum \sigma^{ij}(\bar{x}_i^{(\nu)} - \bar{x}_i)(x_j^{(\nu)} - \bar{x}_j)$ to detect the extreme deviate observation in a multivariate normal sample $x^{(\nu)} = (x_1^{(\nu)}, x_2^{(\nu)}, \dots, x_n^{(\nu)})$, $(\nu = 1, 2, \dots, k)$ assuming that the variance-covariance matrix is known to us, and we shall prove that this test has some optimum property as a multiple decision procedure similar to that of univariate case. (c.f. KUDÔ [3] [4] and PAULSON [5])

It is the opinion of the author that we are justified by the theorem in this paper to call the observation $x^{(\nu)}$ which has the maximum value of $\sum \sigma^{ij}(x_i^{(\nu)} - \bar{x}_i)(x_j^{(\nu)} - \bar{x}_j)$ the extreme value in this multivariate sample, when the variances and the covariances are known to us.

We are further tempted to assume that we are justified to call the observation $x^{(\nu)}$ which has the maximum value of $\sum s^{ij}(x_i^{(\nu)} - \bar{x}_i)(x_j^{(\nu)} - \bar{x}_j)$ the extreme value, where s_{ij} is the estimate of the σ_{ij} obtained from the same sample, and (s^{ij}) is the inverse matrix of (s_{ij}) . For the justification of this assumption, we have to

generalize the theorem to the case when (σ_{ij}) is unknown to us, which the author has not succeeded and wishes to develop on another occasion.

In the section 2 we shall give the formulation of the problem and the statement of the main theorem, whose proof will be given in the section 3. The section 4 is devoted to some generalizations, which will be of some theoretical interest.

The author would like to express his hearty thanks to Professor T. KITAGAWA for the encouragement and criticism given to him while this paper was being prepared.

§ 2. The formulation of the problem and the main theorem

Let $(x_1^{(\nu)}, x_2^{(\nu)}, \dots, x_k^{(\nu)}) = x^{(\nu)}$, $(\nu = 1, 2, \dots, n)$ be distributed in a multivariate normal distribution with mean vector $m^{(\nu)} = (m_1^{(\nu)}, m_2^{(\nu)}, \dots, m_k^{(\nu)})$, $(\nu = 1, 2, \dots, n)$ respectively and a common variance-covariance matrix (σ_{ij}) . We assume that they are mutually independent and the values of the means are unknown to us whereas the values of the elements of the variance-covariance matrix are known to us.

Our null hypothesis H_0 is $H_0 = H(m^{(1)} = m^{(2)} = \dots = m^{(n)} = m)$ where m is free and we have n alternative hypothesis $H_1, H_2, \dots, H_n$ where $H_\nu (\nu = 1, 2, \dots, n)$ is the hypothesis $H_\nu(\Delta m) = H(m^{(1)} = \dots = m^{(\nu-1)} = m^{(\nu)} - \Delta m = m^{(\nu+1)} = \dots = m^{(n)})$ where Δm is a non-zero vector. In other words, under H_0 all the k variate observations are assumed to be all from the same population whereas under the alternative hypothesis H_ν the ν -th observation $x^{(\nu)}$ only is assumed to be from a normal population with a mean vector different from those of others and with the same variance-covariance matrix.

Let $D_\nu (\nu = 0, 1, 2, \dots, n)$ be the decision to accept $H_\nu (\nu = 0, 1, 2, \dots, n)$. Our problem in this paper is to find out an optimum decision procedure as to these $n+1$ decisions on the basis of the sample $x = (x^{(1)}, x^{(2)}, \dots, x^{(n)})$ at hand.

It should be noted here that if the variance-covariance matrix is unknown to us this problem is a multivariate analogue of slippage problem treated by E. PAULSON [5] and the author [3] [4].

At first we have to make clear the meaning of optimum in this generalized formulation.

We introduce here the following requirements or criteria (1°)–(4°).

- (1°) We want to accept the decision D_0 with a pre-assigned probability, say, $1-p$ when H_0 is true.
- (2°) The decision procedure must be invariant under the addition of any constant vector to all the observations in a general sense. (c.f. KUDÔ [4] Definition

2.2) This condition states the following. Let $d_\nu(x) = d_\nu(x^{(1)}, x^{(2)}, \dots, x^{(n)})$, $(\nu=0, 1, \dots, n)$ be the decision function of our decision procedure, where $d_\nu(x)$, $(x=0, 1, \dots, n)$ denotes the probability of accepting the decision D_ν after observing $x=(x^{(1)}, \dots, x^{(n)})$. Naturally we must have $0 \leq d_\nu(x) \leq 1$, and $\sum_{\nu=0}^n d_\nu(x) = 1$ and to meet our mathematical necessity we have to assume $d_\nu(x)$, $(\nu=0, 1, \dots, n)$ be measurable functions. Then our condition states that $d_\nu(x^{(1)}, \dots, x^{(n)}) = d_\nu(x^{(1)} + y, \dots, x^{(n)} + y)$, $(\nu=0, 1, \dots, n)$ for any constant vector y except possibly for a set of measure zero.

Now we introduce here a quantity $\|\Delta m\|_\sigma = \sum \sigma^{ij} \Delta m_i \Delta m_j$ where (σ^{ij}) is the inverse matrix of the variance-covariance matrix. (σ_{ij}) . It is clear that $\|\Delta m\|_\sigma$ is non-negative.

(3°) The probability of making correct decision D_ν when H_ν is correct $P(D_\nu | H_\nu(\Delta m))$ should be a function of $\|\Delta m\|_\sigma$ only and be independent of ν .

(4°) Under the restrictions of (1°), (2°) and (3°), we want to maximize the probability of making a correct decision D_ν when H_ν is correct.

Now our problem is to find out a decision procedure satisfying the conditions (1°)–(4°). The answer to this problem is given by the following.

THEOREM 1. The decision procedure as to these $n+1$ decisions stated above satisfying the conditions (1°)–(4°) is given by the following

$$(2.1) \quad \begin{aligned} \text{if } \sum_{i,j=1}^k \sigma^{ij} (x_i^{(M)} - \bar{x}_i) (x_j^{(M)} - \bar{x}_j) &\geq \lambda_p, & \text{accept } D_M \\ \text{if } \sum_{i,j=1}^k \sigma^{ij} (x_i^{(M)} - \bar{x}_i) (x_j^{(M)} - \bar{x}_j) &< \lambda_p, & \text{accept } D_0 \end{aligned}$$

where M is the suffix of the observation which attains the maximum of

$$(2.2) \quad \sum_{i,j=1}^k \sigma^{ij} (x_i^{(\nu)} - \bar{x}_i) (x_j^{(\nu)} - \bar{x}_j) \text{ namely } M \text{ is defined by} \\ \sum_{i,j=1}^k \sigma^{ij} (x_i^{(M)} - \bar{x}_i) (x_j^{(M)} - \bar{x}_j) = \text{Max}_{\nu=1, \dots, n} \sum_{i,j=1}^k \sigma^{ij} (x_i^{(\nu)} - \bar{x}_i) (x_j^{(\nu)} - \bar{x}_j)$$

and λ_p be so chosen that the probability of correct decision D_0 when H_0 is correct $P(D_0 | H_0)$ is equal to $1-p$.

§ 3. Proof of the theorem

At first we need the following

LEMMA. 3.1. The necessary and sufficient condition for the decision func-

tion $d_\nu(x^{(1)}, \dots, x^{(n)})$, $(\nu=0, 1, \dots, n)$ to satisfy condition (2°) is that

$$(3.1) \quad d_\nu(x^{(1)}, \dots, x^{(n)}) = d_\nu(x^{(1)} - x^{(n)}, \dots, x^{(n-1)} - x^{(n)}, 0), (\nu=0, 1, \dots, n)$$

holds true except possibly for a set of measure zero.

PROOF. The proof is omitted as it is a simple generalization of the Theorem given in a previous paper of the author [4]. (c.f. Theorem 2.1)

As we can restrict our attention to the class of decision procedures which depends only on $x^{(1)} - x^{(n)}, x^{(2)} - x^{(n)}, \dots, x^{(n-1)} - x^{(n)}$ their joint distribution will be of our interest. We do not need, however, the joint frequency function itself, but a function related to it.

Let $y^{(\nu)} = (y_1^{(\nu)}, y_2^{(\nu)}, \dots, y_k^{(\nu)})$, $(\nu=1, 2, \dots, n-1)$ be defined by

$$(3.2) \quad \begin{aligned} y^{(\nu)} &= x^{(\nu)} - x^{(n)}, & (\nu=1, 2, \dots, n-1) \\ y_i^{(\nu)} &= x_i^{(\nu)} - x_i^{(n)}, & (i=1, 2, \dots, k) \end{aligned}$$

and $g(y^{(1)}, y^{(2)}, \dots, y^{(n-1)})$ be their joint frequency function,

and let $h(x^{(1)}, x^{(2)}, \dots, x^{(n)})$ be defined by

$$(3.3) \quad h(x^{(1)}, x^{(2)}, \dots, x^{(n)}) = g(x^{(1)} - x^{(n)}, x^{(2)} - x^{(n)}, \dots, x^{(n-1)} - x^{(n)}).$$

LEMMA 3.2. Let $x^{(\nu)}$, $(\nu=1, 2, \dots, n)$ be independently distributed in multivariate normal distributions with mean vectors $m^{(\nu)}$, $(\nu=1, 2, \dots, n)$ respectively and with a common variance-covariance matrix $A = (\sigma_{ij})$, then the function defined by (3.3) is given by

$$(3.4) \quad \begin{aligned} h(x^{(1)}, x^{(2)}, \dots, x^{(n)}) &= \frac{1}{(\sqrt{2\pi})^{(n-1)k}} \frac{1}{n^k |A|^{n-1}} \exp \left[-\frac{1}{2n} Q_1 \right] \\ Q_1 &= \sum_{i,j=1}^k \sigma^{ij} \left\{ \left(n \sum_{\nu=1}^n x_i^{(\nu)} x_j^{(\nu)} - \left(\sum_{\nu=1}^n x_i^{(\nu)} \right) \left(\sum_{\nu=1}^n x_j^{(\nu)} \right) \right) \right. \\ &\quad - 2 \left(n \sum_{\nu=1}^n x_i^{(\nu)} m_j^{(\nu)} - \left(\sum_{\nu=1}^n x_i^{(\nu)} \right) \left(\sum_{\nu=1}^n m_j^{(\nu)} \right) \right) \\ &\quad \left. + \left(n \sum_{\nu=1}^n m_i^{(\nu)} m_j^{(\nu)} - \left(\sum_{\nu=1}^n m_i^{(\nu)} \right) \left(\sum_{\nu=1}^n m_j^{(\nu)} \right) \right) \right\} \end{aligned}$$

where (σ^{ij}) is the inverse matrix of $A = (\sigma_{ij})$.

PROOF We can easily verify that the variance-covariance matrix of these $k(n-1)$ variables $y_1^{(1)}, y_1^{(2)}, \dots, y_1^{(n-1)}, y_2^{(1)}, y_2^{(2)}, \dots, y_2^{(n-1)}, \dots, y_k^{(1)}, y_k^{(2)}, \dots, y_k^{(n-1)}$ is given by $(\sigma_{ij}) \times A$, namely the Kronecker product of (σ_{ij}) and A where $A = (a_{ij})$ is defined by $a_{ij} = 1 + \delta_{ij}$; $\delta_{ii} = 1$, $\delta_{ij} = 0 (i \neq j)$

and

$$(3.5) \quad \begin{pmatrix} 2\sigma_{11} & \sigma_{11} \cdots \sigma_{11} & 2\sigma_{12} & \sigma_{12} \cdots \sigma_{12} \cdots 2\sigma_{1k} & \sigma_{1k} \cdots \sigma_{1k} \\ \sigma_{11} & 2\sigma_{11} \cdots \sigma_{11} & \sigma_{12} & 2\sigma_{12} \cdots \sigma_{12} \cdots \sigma_{1k} & 2\sigma_{1k} \cdots \sigma_{1k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sigma_{11} & \sigma_{11} \cdots 2\sigma_{11} & \sigma_{12} & \sigma_{12} \cdots 2\sigma_{12} \cdots \sigma_{1k} & \sigma_{1k} \cdots 2\sigma_{1k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 2\sigma_{1k} & \sigma_{1k} \cdots \sigma_{1k} & 2\sigma_{2k} & \sigma_{2k} \cdots 2\sigma_{2k} \cdots 2\sigma_{kk} & \sigma_{kk} \cdots \sigma_{kk} \\ \sigma_{1k} & 2\sigma_{1k} \cdots \sigma_{1k} & \sigma_{2k} & 2\sigma_{2k} \cdots \sigma_{2k} \cdots \sigma_{kk} & 2\sigma_{kk} \cdots \sigma_{kk} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sigma_{1k} & \sigma_{1k} \cdots 2\sigma_{1k} & \sigma_{2k} & \sigma_{2k} \cdots 2\sigma_{2k} \cdots \sigma_{kk} & \sigma_{kk} \cdots 2\sigma_{kk} \end{pmatrix}$$

As the inverse of the matrix A is given by $A^{-1} = (a^{ij})$, $a^{ij} = -1 + n\delta_{ij}$ the inverse of the variance-covariance matrix in our concern is given by

$$(A \times A)^{-1} = A^{-1} \times A^{-1} = (\sigma^{ij}) \times (a^{ij})$$

and the joint density function of $y_i^{(\nu)}$'s is given by

$$(3.6) \quad g(y^{(1)}, y^{(2)}, \dots, y^{(n-1)}) = \frac{1}{(\sqrt{2\pi})^{(n-1)k}} \frac{1}{n^k |A|^{n-1}} \exp \left[-\frac{1}{2n} Q_2 \right]$$

$$Q_2 = (n-1) \sum_{i=1}^k \sigma^{ii} \sum_{\nu=1}^{n-1} (y_i^{(\nu)} - M_i^{(\nu)})^2$$

$$- \sum_{i=1}^k \sigma^{ii} \left(\sum_{\nu=1}^{n-1} y_i^{(\nu)} - M_i^{(\nu)} \right)^2$$

$$+ (n-1) \sum_{i \neq j} \sigma^{ij} \sum_{\nu=1}^{n-1} (y_i^{(\nu)} - M_i^{(\nu)}) (y_j^{(\nu)} - M_j^{(\nu)})$$

$$- \sum_{i \neq j} \sigma^{ij} \left(\sum_{\nu=1}^{n-1} y_i^{(\nu)} - M_i^{(\nu)} \right) \left(\sum_{\nu=1}^{n-1} y_j^{(\nu)} - M_j^{(\nu)} \right)$$

where $M_i^{(\nu)} = m_i^{(\nu)} - m_i^{(n)}$, $(\nu = 1, 2, \dots, n-1, i = 1, 2, \dots, k)$, and Q_2 is equal to

$$(3.7) \quad Q_3 = n \sum_{\nu=1}^{n-1} \sum_{i,j} \sigma^{ij} (y_i^{(\nu)} - M_i^{(\nu)}) (y_j^{(\nu)} - M_j^{(\nu)})$$

$$- \sum_{i,j} \sigma^{ij} \left(\sum_{\nu=1}^{n-1} y_i^{(\nu)} - M_i^{(\nu)} \right) \left(\sum_{\nu=1}^{n-1} y_j^{(\nu)} - M_j^{(\nu)} \right).$$

After inserting the relation, $y_i^{(\nu)} = x_i^{(\nu)} - x_i^{(n)}$ and $M_i^{(\nu)} = m_i^{(\nu)} = m_i^{(\nu)} - m_i^{(n)}$, we have

that the first term in Q_3 is given by

$$\begin{aligned}
 & n \sum_{\nu=1}^n \sum_{i,j} \sigma^{ij} (x_i^{(\nu)} - m_i^{(\nu)}) (x_j^{(\nu)} - m_j^{(\nu)}) \\
 (3.8) \quad & - 2n \sum_{\nu=1}^n \sum_{i,j} \sigma^{ij} (x_i^{(\nu)} - m_i^{(\nu)}) (x_j^{(\nu)} - m_j^{(n)}) \\
 & + n^2 \sum_{i,j} \sigma^{ij} (x_i^{(n)} - m_i^{(n)}) (x_j^{(n)} - m_j^{(n)})
 \end{aligned}$$

and the second one is

$$\begin{aligned}
 & \sum_{i,j} \sigma^{ij} \left(\sum_{\nu=1}^n x_i^{(\nu)} - m_i^{(\nu)} \right) \left(\sum_{\nu=1}^n x_j^{(\nu)} - m_j^{(\nu)} \right) \\
 (3.9) \quad & - 2n \sum_{\nu=1}^n \sum_{i,j} \sigma^{ij} (x_i^{(\nu)} - m_i^{(\nu)}) (x_j^{(n)} - m_j^{(n)}) \\
 & + n^2 \sum_{i,j} \sigma^{ij} (x_i^{(n)} - m_i^{(n)}) (x_j^{(n)} - m_j^{(n)}).
 \end{aligned}$$

And finally we have that it is expressed in terms of $x^{(1)}, x^{(2)}, \dots, x^{(n)}$ and $m^{(1)}, m^{(2)}, \dots, m^{(n)}$ as is given in this Lemma.

COROLLARY. (i) In case H_0 is correct, $h(x)$ becomes

$$(3.10) \quad h_0(x) = \frac{1}{(\sqrt{2\pi})^{(n-1)k}} \frac{1}{n^k |A|^{n-1}} \exp \left[-\frac{1}{2} \sum_{i,j} \sigma^{ij} S_{ij} \right].$$

(ii) In case $H_\nu(\Delta m)$ is correct, $h(x)$ becomes

$$\begin{aligned}
 (3.11) \quad h_\nu(x) &= \frac{1}{(\sqrt{2\pi})^{(n-1)k}} \frac{1}{n^k |A|^{n-1}} \exp \left[-\frac{1}{2} \sum_{i,j} \sigma_{ij} S_{ij} \right] \\
 &\quad \exp \left[-\frac{n-1}{2n^2} \sum_{i,j} \sigma^{ij} (\Delta x_i^{(\nu)} - \Delta m_i) (\Delta x_j^{(\nu)} - \Delta m_j) \right]
 \end{aligned}$$

where

$$\begin{aligned}
 (3.12) \quad S_{ij} &= \frac{1}{n} \sum_{\nu=1}^n (x_i^{(\nu)} - \bar{x}_i) (x_j^{(\nu)} - \bar{x}_j) \\
 \Delta x_i^{(\nu)} &= x_i^{(\nu)} - \bar{x}_i \\
 \bar{x}_i &= \frac{1}{n} \sum_{\nu=1}^n x_i^{(\nu)}.
 \end{aligned}$$

We shall here consider the integral of the form

$$(3.13) \quad \int A d_0(x) f_0(x, m) + \sum_{\nu=1}^n d_\nu(x) \int f_\nu(x; m, \Delta m) d\mu_c(\Delta m) dx$$

where A is some positive constant, $(d_0(x), d_1(x), \dots, d_n(x))$ is a decision function

as to $n+1$ decisions in our concern, $f_0(x, m)$ is the joint frequency function of x satisfying the hypothesis H_0 , $f_\nu(x, m, \Delta m)$, is the same satisfying the hypothesis H_ν with the mean vector of the ν -th observation different from that of others, m , by the amount of Δm , and $\mu_c(\Delta m)$ is an apriori probability distribution with

$$(3.14) \quad \mu_c(\Delta m; \sum_{i,j} \sigma^{ij} \Delta m_i \Delta m_j = \|\Delta m\|_\sigma = c) = 1.$$

The value of this integral depends on A , $\mathcal{A}$, m , Δm the apriori probability distribution μ_c and the decision function $(d_0(x), d_1(x) \dots, d_n(x))$. If we restrict ourselves to the class of decision functions satisfying the condition 2° , this integral does not depend on m .

We shall here prove the following

LEMMA 3.3. For any λp , $\mathcal{A}$ and c , we can find out $\mathcal{A}$ and μ_c such that the decision function of the decision procedure defined by (2.1) is the one which maximizes the integral (3.13) among all the decision procedures satisfying the condition (2°) .

PROOF. At first we shall consider the problem as to maximizing the integral class of when A , $\mathcal{A}$, Δm and μ_c are all fixed. As we restrict our attention to the invariant decision procedures in the sense of condition (2°) , we appeal to the method adopted in the Theorem 2.2. of the author's previous paper [4].

Let $d_\nu(y)$ ($\nu=0, 1, \dots, n$) be defined by

$$(3.18) \quad d_\nu(y) = d_\nu(y^{(1)}, \dots, y^{(n-1)}, 0)$$

where

$$(3.19) \quad y^{(\nu)} = x^{(\nu)} - x^{(n)}$$

In virtue of Lemma 1 the maximizing the integral (3.13) under the condition (2°) is equivalent to maximizing the integral (3.20) without any condition,

$$(3.20) \quad \int A d_0(y) g_0(y) + \sum_{\nu=1}^n d_\nu(y) \int g_\nu(y; \Delta m) d\mu_c(\Delta m) dy$$

where $g_0(y)$ and $g_\nu(y, \Delta m)$, ($\nu=1, 2, \dots, n$) be the joint density function of y under the hypothesis H_0 and H_ν ($\nu=1, 2, \dots, n$) with the amount of slippage Δm respectively. A slight modification of the Theorem 5.1. of WALD [6] gives us that the integral (3.20) is maximized if

$$(3.21) \quad d_0(y) = 1 \text{ if } A g_0(y) \geq \int g_\nu(y, \Delta m) d\mu_c(\Delta m), (\nu=1, 2, \dots, n)$$

and

$$\begin{aligned}
(3.22) \quad & d_\nu(y) = 1 \text{ if} \\
& Ag_0(y) < \int g_\nu(y, \Delta m) d\mu_c(\Delta m) \\
& \int g_\mu(y, \Delta m) d\mu_c(\Delta m) < \int g_\nu(y, \Delta m) d\mu_c(\Delta m), (\mu = 1, 2, \dots, n, \mu \neq \nu)
\end{aligned}$$

After inverting this relation in terms of x , we have that the integral is maximized under the restriction (2°), if $d_0(x)$. $d_1(x), \dots, d_n(x)$ are defined by the following conditions

$$(3.23) \quad d_0(x) = 1 \text{ if } Ah_0(x) > \text{Max} \int h_\nu(x, \Delta m) d\mu_c(\Delta m), (\nu = 1, 2, \dots, n)$$

and

$$\begin{aligned}
(3.24) \quad & d_\nu(x) = 1 \text{ if} \\
& Ah_0(x) \leq \int h_\nu(x, \Delta m) d\mu_c(\Delta m) \\
& \int h_\mu(x, \Delta m) d\mu_c(\Delta m) \leq \int h_\nu(x, \Delta m) d\mu_c(\Delta m), (\mu = 1, 2, \dots, n, \mu \neq \nu).
\end{aligned}$$

For any C , we shall define $\mu_c(\Delta m)$ such that if the surface of the ellipsoid defined by

$$(3.25) \quad \sum_{i,j} \sigma^{ij} \Delta m_i \Delta m_j = C$$

is transformed to the surface of the unit k -dimensional sphere by a linear transformation, the measure induced on this surface by this linear transformation from the measure μ_c is an uniform distribution on it. Let $S(\Delta m)$ denote the uniform distribution on the unit k -dimensional sphere. Then we have after making the linear transformation such as above

$$\begin{aligned}
(3.26) \quad & \int h_\nu(x, \Delta m) d\mu_c(\Delta m) \\
& = \frac{1}{(1/2\pi)^{(n-1)k} n^k |\sigma_{ij}|^{n-1}} \exp\left(-\frac{1}{2} \sum_{i,j} \sigma^{ij} S_{ij}\right) \exp\left(-\frac{n-1}{2n^2} \sum_{i,j} \sigma^{ij} \Delta x_i^{(\nu)} \Delta x_j^{(\nu)}\right) \\
& \quad \int \exp\left\{\frac{n-1}{2n^2} \sum_{i,j} \sigma^{ij} (\Delta x_i^{(\nu)} - \Delta m_i) (\Delta x_j^{(\nu)} - \Delta m_j)\right\} d\mu_c(\Delta m) \\
& = C \exp\left[-\frac{1}{2} \sum_{i,j} \sigma^{ij} S_{ij}\right] \exp\left[+\frac{n-1}{2n^2} \sum_{i,j} \sigma^{ij} \Delta x_i^{(\nu)} \Delta x_j^{(\nu)}\right] \\
& \quad \int \exp\left[\frac{n-1}{2n^2} \sum_i (\Delta m_i - y_i)^2\right] dS(\Delta m)
\end{aligned}$$

where C is some constant common for all k , and $\sum_{i=1}^k y_i = \sum_{i,j=1}^k \sigma^{ij} \Delta x_i^{(\nu)} \Delta x_j^{(\nu)}$.

As $S(\Delta m)$ is a uniform distribution we have

$$\begin{aligned}
& \int h_\nu(x, \Delta m) d\mu_c(\Delta m) \\
&= C \exp \left[-\frac{1}{2} \sum_{i,j} \sigma^{ij} S_{ij} \right] \exp \left[+\frac{n-1}{2n^2} \sum_{i,j} \sigma^{ij} \Delta x_i^{(\nu)} \Delta x_j^{(\nu)} \right] \\
& \quad \int \exp \left[\frac{n-1}{2n^2} \left\{ \sum_{j=1}^{k-1} \Delta m_j^2 + (\Delta m_k - \sqrt{\sum_{i,j} \sigma^{ij} \Delta x_i^{(\nu)} \Delta x_j^{(\nu)}}) \right\} \right] dS(\Delta m) \\
(3.27) \quad &= C \exp \left[-\frac{1}{2} \sum_{i,j} \sigma^{ij} S_{ij} \right] \exp \left[+\frac{n-1}{2n^2} \sum_{i,j} \Delta m_i^1 \right] \\
& \quad \int_{\Delta m_k > 0} \exp \left[\frac{n-1}{n^2} \Delta m_k \sqrt{\sum_{i,j} \sigma^{ij} \Delta x_i^{(\nu)} \Delta x_j^{(\nu)}} \right] \\
& \quad + \exp \left[-\frac{n-1}{n^2} \Delta m_k \sqrt{\sum_{i,j} \sigma^{ij} \Delta x_i^{(\nu)} \Delta x_j^{(\nu)}} \right] dS(\Delta m).
\end{aligned}$$

This shows that

$$(3.28) \quad \int h_\nu(x, \Delta m) d\mu_c(\Delta m) \geq \int h_\mu(x, \Delta m) d\mu_c(\Delta m)$$

if and only if

$$(3.29) \quad \sum_{i,j} \sigma^{ij} \Delta x_i^{(\nu)} \Delta x_j^{(\nu)} \geq \sum_{i,j} \sigma^{ij} \Delta x_i^{(\mu)} \Delta x_j^{(\mu)}$$

and

$$(3.30) \quad Ah_0(x) \geq \int h_\nu(x, \Delta m) d\mu_c(\Delta m)$$

if and only if

$$(3.31) \quad B \geq \sum_{i,j} \sigma^{ij} \Delta x_i^{(\nu)} \Delta x_j^{(\nu)}$$

where B is some constant suitably chosen. As the right hand side of (3.30) is a strictly monoton increasing function of $\sum \sigma^{ij} \Delta x_i^{(\nu)} \Delta x_j^{(\nu)}$ there exists a positive constant A such that

$$(3.32) \quad Ah_0(x) \geq \text{Max}_{\nu=1,2,\dots,n} \int h_\nu(x, \Delta m) d\mu_c(\Delta m)$$

if and only if

$$(3.33) \quad \lambda_p \geq \text{Max}_{\nu=1,\dots,n} \sum_{i,j} \sigma^{ij} \Delta x_i^{(\nu)} \Delta x_j^{(\nu)}.$$

These two facts above lead us to the conclusion.

PROOF OF THE THEOREM 1. Let $(\bar{d}_0(x), \dots, \bar{d}_n(x)) = \bar{d}(x)$ be the decision function of the decision procedure defined by (2.1), then we have to prove that there does not exist another decision function $(\bar{d}_0(x), \bar{d}_1(x), \dots, \bar{d}_n(x)) = \bar{\bar{d}}(x)$ such that $\bar{\bar{d}}(x)$ is better than $\bar{d}(x)$ at some parameter point m and Δm .

That is we have

$$(3.34) \quad \int \bar{d}_0(x) f_0(x, m) dx = \int \bar{d}_0(x) f_0(x, m) dx = 1 - p,$$

and

$$(3.35) \quad \int \bar{d}_\nu(x) f_\nu(x, m, \Delta m) dx \leq \int \bar{d}_\nu(x) f_\nu(x, m, \Delta m) dx, (\nu = 1, 2, \dots, n).$$

If such a decision function would exist, we should have

$$(3.36) \quad \begin{aligned} & \int A \bar{d}_0(x) f_0(x, m) + \sum_{\nu=1}^n \bar{d}_\nu(x) \int f_\nu(x, m, \Delta m) d\mu_c(\Delta m) dx \\ & \leq \int A \bar{d}_0(x) f_0(x, m) + \sum_{\nu=1}^n \bar{d}_\nu(x) \int f_\nu(x, m, \Delta m) d\mu_c(\Delta m) dx \end{aligned}$$

for any positive constant A and for any μ_c such that $\mu_c\{\Delta m; \sum \sigma^{ij} \Delta m_i; \Delta m_j = C\} = 1$, which leads us to the contradiction to Lemma 3. Q.E.D.

§4. Some generalizations

As is indicated in the section 1, the assumption that the a variance-covariance matrix is known to us appears to be too restrictive in practice. We can replace this assumption by a more general one that $(x_1^{(\nu)}, x_2^{(\nu)}, \dots, x_k^{(\nu)})$, $(\nu = 1, 2, \dots, n)$ have a same variance-covariance matrix $(\sigma_{ij}) = \sigma^2(\rho_{ij})$ where (ρ_{ij}) is known whereas σ^2 is unknown to us.

Under this assumption, we shall consider the same multiple decision problem as to the $n+1$ decisions $D_0, D_1, \dots, D_n$ as defined in the section 2.

In this case we have to replace conditions (2°) and (3°) by $(2^{\circ\circ})$ and $(3^{\circ\circ})$ as follows.

$(2^{\circ\circ})$ The decision procedure must be invariant under the addition of any constant vector and also under the multiplication of any constant scalar to all the observation in a general sense. (c.f. KUDÔ [4] Definition 2.2)

This condition is equivalent to the one that the decision function of our decision procedure should satisfy the following

$$(4.1) \quad d_\nu(x^{(1)}, x^{(2)}, \dots, x^{(n)}) = d_\nu(\alpha x^{(1)} + y, \alpha x^{(2)} + y, \dots, \alpha x^{(n)} + y), (\nu = 0, 1, \dots, n)$$

except possibly for a set of measure zero where α is any nonzero constant and y is any constant vector.

$(3^{\circ\circ})$ The possibility of making correct decision D_ν when H_ν is correct $P_\nu(D_\nu | H_\nu)$ should be a function of $\|\Delta m\|_\sigma / \sigma^2$ only and be independent of ν .

The answer to the decision problem in this case is given by the following

THEOREM 2. The decision procedure as to these $n+1$ decisions stated above satisfying the conditions (1°) , $(2^{\circ\circ})$, $(3^{\circ\circ})$ and (4°) is given by the following

$$(4.2) \quad \begin{aligned} & \text{if } \sum_{i,j=1}^k \rho^{ij} (x_i^{(M)} - \bar{x}_i)(x_j^{(M)} - \bar{x}_j) \geq \mu_p \left| \sum_{i,j=1}^k \rho^{ij} s_{ij} \right|, \text{ accept } D_M \\ & \text{if } \sum_{i,j=1}^k \rho^{ij} (x_i^{(M)} - \bar{x}_i)(x_j^{(M)} - \bar{x}_j) < \mu_p \left| \sum_{i,j=1}^k \rho^{ij} s_{ij} \right|, \text{ accept } D_0 \end{aligned}$$

where M is the suffix of the observation which attains the maximum of $\sum_{i,j} \rho^{ij} (x_i^{(\gamma)} - \bar{x}_i)(x_j^{(\gamma)} - \bar{x}_j)$ namely M is defined by

$$(4.3) \quad \sum_{i,j=1}^k \rho^{ij} (x_i^{(M)} - \bar{x}_i)(x_j^{(M)} - \bar{x}_j) = \text{Max}_{\gamma=1,2,\dots,n} \sum_{i,j=1}^k \rho^{ij} (x_i^{(\gamma)} - \bar{x}_i)(x_j^{(\gamma)} - \bar{x}_j)$$

and (ρ^{ij}) is the inverse matrix of (ρ_{ij}) and μ_p be so chosen that the probability of making correct decision D_0 when H_0 is correct $P(D_0|H_0)$ is equal to $1-p$.

PROOF. We shall briefly state the proof of this theorem.

Instead of Lemma 3, we need the following

LEMMA 4.1. The necessary and sufficient condition for the decision function $d_\nu(x^{(1)}, x^{(2)}, \dots, x^{(n)})$, $(\nu=0, 1, 2, \dots, n)$ to satisfy the condition $(2^{\circ\circ})$ is that

$$(4.4) \quad d_\nu(x^{(1)}, \dots, x^{(n)}) = d_\nu\left(\frac{x^{(1)} - x^{(n)}}{x_k^{(n-1)} - x_k^{(n)}}, \dots, \frac{x^{(n-1)} - x^{(n)}}{x_k^{(n-1)} - x_k^{(n)}}, 0, \dots, 0\right), (\nu=0, 1, \dots, n)$$

holds true except for a set of measure zero.

PROOF The proof is given by the combination of Lemma 3.1. of this paper and the Theorem 2.1. of the author's previous paper. (KUDÔ [4]) Because of the Lemma above, we are here interested in the density function of $(z_1^{(\gamma)}, z_2^{(\gamma)}, \dots, z_k^{(\gamma)})$, $(\nu=1, 2, \dots, n-1)$ defined by

$$(4.5) \quad \begin{aligned} z^{(\nu)} &= y^{(\nu)} / y_k^{(n-1)} = (x^{(\nu)} - x^{(n)}) / (x_k^{(n-1)} - x_k^{(n)}) \\ z_i^{(\nu)} &= y_i^{(\nu)} / y_k^{(n-1)} = (x_i^{(\nu)} - x_i^{(n)}) / (x_k^{(n-1)} - x_k^{(n)}) \end{aligned}$$

It may be noted that $z_k^{(n-1)}$ is not a random variable but a constant equal to 1. Let $g'(z^{(1)}, \dots, z^{(n-1)})$ be their joint density function and let $h'(x^{(1)}, \dots, x^{(n)})$ be defined by

$$(4.6) \quad h'(x^{(1)}, \dots, x^{(n)}) = g'\left(\frac{x^{(1)} - x^{(n)}}{x_k^{(n-1)} - x_k^{(n)}}, \dots, \frac{x^{(n-1)} - x^{(n)}}{x_k^{(n-1)} - x_k^{(n)}}\right)$$

Then we have corresponding to the Lemma 3.2. the following

LEMMA 4.2. Under the same assumption as to that of Lemma 3.2, we have

$$\begin{aligned}
h'(x^{(1)}, x^{(2)}, \dots, x^{(n)}) &= \frac{1}{(\sqrt{2\pi})^{(n-1)k} n^k |A|^{n-1}} |x_k^{(n-1)} - x_k^{(n)}|^{-(n-1)k} \\
&\quad \int_{-\infty}^{+\infty} |t|^{(n-1)k-1} \exp\left(-\frac{1}{2n} Q_4\right) dt \\
(4.7) \quad Q_4 &= \sum_{i,j=1}^k \sigma^{ij} \left\{ t^2 \left(n \sum_{\nu=1}^n x_i^{(\nu)} x_j^{(\nu)} - \left(\sum_{\nu=1}^n x_i^{(\nu)} \right) \left(\sum_{\nu=1}^n x_j^{(\nu)} \right) \right) \right. \\
&\quad - 2t \left(n \sum_{\nu=1}^n x_i^{(\nu)} m_j^{(\nu)} - \left(\sum_{\nu=1}^n x_i^{(\nu)} \right) \left(\sum_{\nu=1}^n m_j^{(\nu)} \right) \right) \\
&\quad \left. + \left(n^2 \sum_{\nu=1}^n m_i^{(\nu)} m_j^{(\nu)} - \left(\sum_{\nu=1}^n m_i^{(\nu)} \right) \left(\sum_{\nu=1}^n m_j^{(\nu)} \right) \right) \right\}
\end{aligned}$$

PROOF After making the transformation from $(y_1^{(1)}, \dots, y_k^{(1)}, \dots, y_1^{(n-1)}, \dots, y_k^{(n-1)})$ to $(z_1^{(1)}, \dots, z_k^{(1)}, \dots, z_1^{(n-1)}, \dots, z_k^{(n-1)}, S)$ by making use of the relation (4.5) and $y^{(n-1)}t = S$, and making integration with S , we have from (3.7), the joint frequency function of z

$$\begin{aligned}
(4.8) \quad g'(z) &= \frac{1}{(\sqrt{2\pi})^{(n-1)k} n^k |A|^{n-1}} \int |t|^{(n-1)k-1} \exp\left(-\frac{1}{2n} Q_5\right) dt \\
Q_5 &= n \sum_{i,j=1}^k \left\{ \sum_{\nu=1}^n (z_i^{(\nu)} t - M_i^{(\nu)}) (z_j^{(\nu)} t - M_j^{(\nu)}) - \left(\sum_{\nu=1}^n z_i^{(\nu)} t - M_i^{(\nu)} \right) \left(\sum_{\nu=1}^n z_j^{(\nu)} t - M_j^{(\nu)} \right) \right\}
\end{aligned}$$

After expressing $g'(z)$ in terms of x by making use of (4.5), making the transformation of the variable again $(x_k^{(n-1)} - x_k^{(n)})t = S$ we obtain the result of this Lemma.

Hereafter we can proceed the proof exactly parallel to that of the Theorem 1. We shall only mention the formula corresponding to (3.19). Instead of (3.19) we have

$$\begin{aligned}
(4.9) \quad &\int h'_\nu(x, \Delta m) d\mu_c(\Delta m) \\
&= C |x_k^{(n-1)} - x_k^{(n)}|^{-(n-1)k} \int_{t=-\infty}^{+\infty} |t|^{(n-1)k-1} \int \exp\left[-\frac{1}{2n} Q_6\right] d\mu_c(\Delta m) dt \\
Q_6 &= nt^2 \sum_{i,j} \sigma^{ij} S_{ij} + \frac{n-1}{n} t^2 \sum_{i,j} \sigma^{ij} \Delta x_i^{(\nu)} \Delta x_j^{(\nu)} - \frac{n-1}{n} \\
&\quad \left\{ \sum_{i=1}^{k-1} \Delta m_i^2 + (\Delta m_k - \sqrt{\sum_{i,j} \sigma^{ij} \Delta x_i^{(\nu)} \Delta x_j^{(\nu)} t^2})^2 \right\}
\end{aligned}$$

and after making the transformation $\sqrt{|\sum_{i,j} \sigma^{ij} S_{ij}|} t = S$ we have

$$\begin{aligned}
(4.10) \quad &\int h'(x, \Delta m) d\mu_c(\Delta m) \\
&= C |x_k^{(n-1)} - x_k^{(n)}|^{-(n-1)k} \left| \sum_{i,j} \sigma^{ij} S_{ij} \right|^{-(n-1)k} \int_{S=-\infty}^{+\infty} \int |t|^{(n-1)k-1} \exp\left[-\frac{1}{2n} Q_7\right]
\end{aligned}$$

$$Q_7 = S^2 - \frac{n-1}{n}C + 2\frac{n-1}{n_j} \Delta m_k \frac{\sqrt{\sum_{i,j} \sigma^{ij} \Delta x_i^{(\nu)} \Delta x_j^{(\nu)}}}{\sqrt{\sum_{i,j} \sigma^{ij} S_{ij}}}$$

where C is some constant common for all ν .

Q. E. D.

Now let us consider another type of generalization.

Let $(x_1^{(\nu, \mu)}, x_2^{(\nu, \mu)}, \dots, x_k^{(\nu, \mu)}), (\nu = 1, 2, \dots, n; \mu = 1, \dots, m)$

be independently distributed in a multivariate normal distribution with the mean vector $(m_1^{(\nu)}, m_2^{(\nu)}, \dots, m_k^{(\nu)}) = m^{(\nu)}$ for $x^{(\nu, \mu)}, (\mu = 1, 2, \dots, m)$ and a common variance-covariance matrix. Instead of all the n alternative hypotheses, we shall here consider the case when we have only the first l alternative hypotheses, $H_1, H_2, H_3, \dots, H_l, (l \leq n)$. The same problem arises here and the problems which have been treated in this paper are the special case when $l = n$ and $m = 1$.

In case when the variance-covariance matrix is unknown to us, we have

THEOREM 3. The decision procedure as to these $l+1$ decisions satisfying conditions (1°), (2°), (3°), and (4°) is given by the following

$$(4.11) \quad \begin{aligned} & \text{if } \sum_{i,j=1}^k \sigma^{ij} (\bar{x}_i^{(M)} - \bar{x}_i) (\bar{x}_j^{(M)} - \bar{x}_j) \geq \lambda' p, \text{ accept } D_M \\ & \text{if } \sum_{i,j=1}^k \sigma^{ij} (\bar{x}_i^{(M)} - \bar{x}_i) (\bar{x}_j^{(M)} - \bar{x}_j) < \lambda' p, \text{ accept } D_0 \end{aligned}$$

where M is the suffix of the mean vector whose estimate attains the maximum of $\sum_{i,j=1}^k \sigma^{ij} (\bar{x}_i^{(\nu)} - \bar{x}_i) (\bar{x}_j^{(\nu)} - \bar{x}_j)$ where ν runs from 1 to l , namely M is defined by

$$(4.12) \quad \sum_{i,j=1}^k \sigma^{ij} (\bar{x}_i^{(M)} - \bar{x}_i) (\bar{x}_j^{(M)} - \bar{x}_j) = \text{Max}_{\nu=1,2,\dots,n} \sum_{i,j=1}^k \sigma^{ij} (\bar{x}_i^{(\nu)} - \bar{x}_i) (\bar{x}_j^{(\nu)} - \bar{x}_j)$$

and

$$(4.13) \quad \begin{aligned} \bar{x}_i^{(\nu)} &= \frac{1}{m} \sum_{\mu=1}^m x_i^{(\nu, \mu)} \\ \bar{x} &= \frac{1}{nm} \sum_{\nu=1}^n \sum_{\mu=1}^m x_i^{(\nu, \mu)} \\ S_{ij} &= \frac{1}{nm} \sum_{\nu=1}^n \sum_{\mu=1}^m (x_i^{(\nu, \mu)} - \bar{x}_i) (\bar{x}_j^{(\nu, \mu)} - \bar{x}_j) \end{aligned}$$

and $\lambda'_p = \lambda'_p(n, m, l, A)$ is so chosen that the probability of making correct decision D_0 when H_0 is correct $P(D_0 | H_0)$ is equal to $1 - p$.

Further, in case when the variance-covariance matrix is of the form $(\sigma_{ij}) = \sigma^2(\rho_{ij})$, where σ^2 is unknown but (ρ_{ij}) is known to us, we have

THEOREM 4. The decision procedure as to these $l+1$ decisions satisfying the conditions (1°), (2°), (3°) and (4°) is given by the following

$$(4.14) \quad \begin{aligned} & \text{if } \sum_{i,j=1}^k \rho^{ij}(\bar{x}_i^{(M)} - \bar{x}_i)(\bar{x}_i^{(M)} - \bar{x}_j) \geq \mu'_p \left| \sum_{i,j=1}^k \rho^{ij} S_{ij} \right|, \text{ accept } D_M \\ & \text{if } \sum_{i,j=1}^k \rho^{ij}(\bar{x}_i^{(M)} - \bar{x}_i)(\bar{x}_j^{(M)} - \bar{x}_j) \leq \mu'_p \left| \sum_{i,j=1}^k \rho^{ij} S_{ij} \right|, \text{ accept } D_0 \end{aligned}$$

where M is the suffix of the mean vector whose estimate attains the maximum of $\sum \rho^{ij}(\bar{x}_i^{(\nu)} - \bar{x}_i)(\bar{x}_j^{(\nu)} - \bar{x}_j)$ where ν runs from 1 to l namely M is defined by

$$(4.15) \quad \sum_{i,j=1}^k \rho^{ij}(\bar{x}_i^{(M)} - \bar{x}_i)(\bar{x}_j^{(M)} - \bar{x}_j) = \text{Max}_{\nu=1,2,\dots,l} \sum_{i,j=1}^k \rho^{ij}(\bar{x}_i^{(\nu)} - \bar{x}_i)(\bar{x}_j^{(\nu)} - \bar{x}_j)$$

and other notations are as defined in the Theorem 3.

PROOF OF THEOREM 3 AND 4. There is no essential difference between the proofs of these Theorems and those of the previous ones, except for uninteresting and cumbersome complications in the notation systems.

REFERENCES.

- 1) GUMBEL, E. J. *Statistical theory of extreme values and some practical applications*. Applied Mathematics Series. No. **33**. National Bureau of Standards. 1954.
- 2) KITAGAWA, T. *Successive process of statistical inferences*. Mem. Fac. Sci. Kyusyu Univ. Ser. A. (Math.) Vol. 5 (1950) No. pp. 139-180.
- 3) KUDÔ, A. *On the testing of outlying observations*. Sankhya, Vol. 17. (1956) pp. 67-76.
- 4) KUDÔ, A. *On the invariant multiple decision procedures*. Bull. Math. Stat. Vol. 6 (1956) pp. 57-68.
- 5) PAULSON, E. *An optimum solution to the k sample slippage problem for the normal distribution*. Ann. Math. Stat. Vol. 23. (1952) pp. 610-616.
- 6) WALD, H. *Statistical Decision Functions*. John Wiley & Sons, Inc., New York, 1950.