Sampling Distributions of Statistics Associated with a Fractile Graphic Method

Kitagawa, Toshio
Kyushu University

https://doi.org/10.5109/12989

出版情報：統計数理研究. 9 (1), pp.10-42, 1960-03. Research Association of Statistical Sciences
バージョン：
権利関係：

SAMPLING DISTRIBUTIONS OF STATISTICS ASSOCIATED WITH A FRACTILE GRAPHIC METHOD

By

Tosio KITAGAWA

(*Received* December 20, 1959)

Introduction

0.1. Fundamental notions in fractile graphic analysis. In one of his lecture notes given in Japan 1958, November, P. C. MAHALANOBIS [1] stated his idea on fractile graphical analysis, where the notions of fractile group, fractile graphs, error area and separation play the fundamental rôles. Since these notes have not been published and their circulation seems to be limited, we shall beginn with a quotation of the original enunciations by P. C. MAHALANOBIS [1], 2.1~5.1.

(*a*) **Fractile groups.** "Suppose each sub-sample consists of n elementary units, each unit being a pair of values of the two variates x and y. Consider the first sub-sample of, say, n sample-units. Rank them in order of ascending values of x. It is now possible to divide the n units into g groups $(1, 2, .., i, ..., g)$, each of equal number, say n'; so that $n=gn'$. These may be called fractile groups." (MAHALANOBIS [1], 2.1~2.2)

(*b*) **Fractile graphs, $G(1)$ and $G(2)$.** "Next calculate the mean value (or median) of n' values of y in each group to give in the first sub-sample the values $y_1', y_2', ..., y_g'$ corresponding to the serial number of the groups $1, 2, ..., i, ..., g$. Take g equi-distant points $(1, 2, ..., i, ..., g)$ on the x-axis to represent the g groups; and plot the corresponding values of $y_1', y_2', ..., y_g'$. Finally, draw straight lines to join each pair of adjoining points y_1' and y_2'; y_2' and y_3'; ..., y_{i-1}' and y_i'; ..., y_{g-1}' and y_g'. This connected chain of lines will be called the fractile graph $G(1)$. Now consider the second sub-sample. It is possible to go through a similar process of ranking the sample units in order of ascending values of x; dividing them after ranking into g groups each of equal number n'; calculating the values of y_1'', $y_2'', ..., y_g''$; and drawing a second fractile graph $G(2)$. We can then have two sub-sample graphs $G(1)$ and $G(2)$, which have equal statistical validity." (MAHALANOBIS [1], 3. ~3.3)

(*c*) **Fractile graph $G(1, 2)$.** "It is also possible to mix together the two sub-samples to form a combined sample, and rank the sample units again; divide into g groups of equal number $(2n')$; calculate the new values of $y_1, y_2, \ldots, y_g$; and draw the fractile graph $G(1,2)$ for the two sub-samples taken together, that is, for the combined sample." (MAHALANOBIS [1], 3.4).

(d) **Error areas.** "It is possible to measure on paper the area bounded by the two sub-sample fractile graphs $G(1)$ and $G(2)$. We shall (semi-intuitively) call this area $a(1,2)$ as the "error" to be associated with the combined fractile graph $G(1,2)$." (MAHALANOBIS [1], 4.1).

(e) **Separation.** "It is further possible to consider a second population from which a pair of interpenetrating sub-samples are drawn, and a second set of fractile graphs, say $G'(1)$, $G'(2)$ and $G'(1,2)$ are constructed in exactly the same way. The area bounded by $G'(1)$ and $G'(2)$ would give the second error area $a'(1,2)$ to be associated with the second pooled graph $G'(1,2)$." (MAHALANOBIS [1], 5.1~5.2)

0.2. The needs of sampling theories for fractile graphic analysis. The fractile graph analysis has a lot of practical applications to statistical analysis of data, as has been shown by various examples of P. C. MAHALANOBIS [1]~[2].

"It is also rich of flexibilities because of minor restrictions on the mutual relation between x and y." (MAHALANOBIS [1], 7.1~7.3) Before claiming a powful or statistical tool as a substitute or even an improved alternative for current statistical analysis, it is, however, required to establish some exact observations on sampling distributions of statistics associated with the notions $(a)\sim(e)$. Otherwise, testings of hypothesis and estimation theorems could not be developed in the standard of the current inference theories. There remain theorefore several fundamental observations still unestabished.

0.3. The Surmises by Mahalanobis. P. C. MAHALANOBIS [1] gave a number of semi-intuitive surmises some of which are experimentally studied. Among others he gave the following ones:

(1°) "The area $a(1,2)$ which has been called the error associated with $G(1,2)$ would decrease statistically in proportion to $n'^{-1/2}$ with increasing size of the sample n' of each group (when g is constant); and would increase in proportion to g, (when n' is kept constant); as a first approximation." (MAHALANOBIS [1], 6.2)

(2°) "The combined fractile graph $G(1,2)$ would tend statistically to lie more and more within the area $a(1,2)$ with increasing values of n' (with g constant)." (MAHALANOBIS [1], 6.3)

(3°) "The number of points of intersection of $G(1)$ and $G(2)$ would tend statistically to be distributed like changes in "runs" of heads and tails in g throws of an unbiassed coin." (MAHALANOBIS [1], 6.4)

(4°) "The error to be associated with the "Separation", $S(1,2)$, to be called, say, E can be found in the usual way from the two error areas, $a(1,2)$ and $a'(1,2)$, associated respectively with the two combined fractile graphs, $G(1,2)$ and $G'(1,2)$, representing respectively the two populations from which the two pairs of sub-samples are drawn. That is, it is possible to take $E=\sqrt{[a^2(1,2)+a'^2(1,2)]}$." (MAHALANOBIS [1], 6.5)

(5°) "To test the significance of the observed separation, it is possible to use the criterion S^2/E^2 which would tend to be distributed, as a first approximation, like Chi-square." (MAHALANOBIS [1], 6.6)

0.4. Summary of the present paper. This paper concerns with the more fundamental aspects of sampling distributions associated with statistics of fractile graphic analysis rather than with the Mahalanobis surmises. Let us start with a random sample $\{(x_i, y_i)\}$ $(i=1,2,\ldots,n)$ of size n from a bivariate population. Let $\{x_{(i)}\}$ be the order statistics defined by $\{x_i\}$ $(i=1,2,\ldots,n)$ such that $x_{(1)}\leqq x_{(2)}\leqq\ldots\leqq x_{(n)}$ which is a rearrangement of n values $x_1, x_2, \ldots, x_n$ in the ascending order. Each y_j associated with x_j in our random sample will be denoted by $y_{((i))}$ when x_j is corresponding to $x_{(i)}$ in the rearrangement. Let us divide the set $\{y_{((j))}\}$ $(j=1,2,\ldots,n)$ into g subsets G_i $(i=1,2,\ldots,g)$ where each G_i consists of $\{y_{((N_{i-1}+k))}\}$ $(k=1,2,\ldots,n_i')$, such that $N_0=0$, $N_{i-1}=n_1'+n_2'+\ldots+n_{i-1}'$ $(i=1,2,\ldots g)$ with $n_1'+n_2'+\ldots+n_g'=n$, and let us define

$$(1.02)\qquad y_{((i))}=(y_{((N_{i-1}+1))}+y_{((N_{i-1}+2))}+\ldots+y_{((N_i))})/n_i',\ (i=1,2,\ldots,g).$$

We are mainly (not exclusively) concerned with the case when $n_1'=n_2'\ldots=n_g'=n'$ and $gn'=n$.)

Let us consider the case when $n_1'=n_2'=\ldots=n_g'=n$ and $gn'=n$. In virtue of the statistics $\{y_{((i))}\}$ we can define the area S which corresponds to the error area in the terminology of MAHALANOBIS [1]. The area S is the sum of $(g-1)$ areas $\{S_k\}$ $(k=1,2,\ldots,g-1)$ where each S_k is defined with respect to the four points (1.11). This definition can be generalised to the separation in the sense of MAHALANOBIS, as we will show in § 1.

The integrals (3.21) and (3.22) are directly associated with the evaluations of $E\{S_k^l\}$ under our particular situations appealing to the asymptotic normality.

The results in § 1 are concerned with the exact representations of (1°) the simultaneous distribution functions of some of the statistics $\{y_{((i))}\}$,(2°) the distribution function of $(y_{((h+1))}+\ldots+y_{((k))})/(k-h)$, (3°) $E\{S_k^l\}$, (4°) $E\{S_{k1}^{l1}\ S_{k2}^{l2}\}$, (5°) $E\{S_{k1}^{l1}\ S_{k2}^{l2}\ S_{k3}^{l3}\}$, (6°) $E\{S_{k1}^{l1}\ S_{k2}^{l2}\ S_{k3}^{l3}\ S_{k4}^{l4}\}$, where $l_i\geqq 1$ and $1\leqq k_1<k_2<k_3<k_4\leqq g$, and hence (7°) $E\{S^l\}$ $(l=1,2,3,4)$.

In § 2 the asymptotic form of the multivariate probability density function of $(\bar{y}_{((k_1))}, \bar{y}_{((k_2))}, \ldots, y_{((k_h))})$ is discussed on the basis of the Assumptions I and II. The Assumption I refers to the expression of the simultaneous probability density functions in the form of (1.32), and will yield us normal approximations to the multivariate probability density functions. The Assumption II is introduced so as to be able for us to appeal to the well-known Theorem of MOSTELLER [1] concerning the limiting simulataneous distribution of order statistics. Here the asymptotic normality is established in Theorem 2.1 for the simultaneous probability distribution of the statistics $(\bar{y}_{((k_1))}, y_{((k_2))}, \ldots, \bar{y}_{((kh))})$ as n' tends to infinity. In view of the results given in Theorem 2.1 and in § 1.8, the evaluations of the integrals enunciated in the right-hand sides of $(1°) \sim (15°)$ in § 1.8 such as those given in $(1.521) \sim (1.524)$ are reduced to those associated with the multivariate normal populations, provided that we are interested with a fairly large or moderately large n'. In § 3.1 we start with a general procedure to evaluate the integrals of the type (3.01) associated with an n-dimentional multivariate normal distribution. In § 3.2 the procedure is applied to the two-dimensional case, that is, $n=2$.

The results given in § 3.2 can directly serve to give asymptotic evaluations of the first two moments of the area S, that is, $E\{S^l\}$ $(l=1, 2)$ under the assumption of asymptotic normality. The general procedure of evaluating the integral of the type (3.01) can also serve to calculate the asymptotic values of the third and the fourth moments $E\{S^l\}$ $(l=3, 4)$. In view of tremendous numbers of various types of the integrals in the right-hand sides of (1.47) and (1.48), however, the problem is remaining unsolved in this paper how to evaluate the sums of them, although their summands can be shown to have certain rather complicated asymptotic values.

In the consequence the surmises by MAHALANOBIS [1] are not claimed to have been solved in any definite way in this paper, but it is hoped that our formulation and allied analysis will serve to make an approach to them. Numerical aspects are shown in § 4 to give certain constants such as variances, covariances and correlation coefficients associated with the multivariate normal distribution derived from a graphical fractile analysis when the sample size is 20.

It is also to be noted that K. TAKEUCHI [1] discussed the surmises by MAHALANOBIS on the basis of regression model which is different from the formulation of the present paper, but had reached the results which are essentially the same with ours given in $(3.461) \sim (3.462)$.

§ 1. The statistics $\{y_{((i))}\}$ and $\{S_k\}$ defined for a bivariate distribution (X, Y) with respect to the order statistics of X.

1.1. Let $\{(x_i, y_i)\}$ $(i=1, 2, \ldots, n)$ be a random sample of size n from a bivariate population. Let $\{x_{(i)}\}$ be the order statistics defined by $\{x_i\}$ $(i=1, 2, \ldots, n)$ such that $x_{(1)} \leqq x_{(2)} \leqq \ldots \leqq x_{(n)}$ which is a rearrangement of n values $x_1, x_2, \ldots, x_n$ in the ascending order. Each y_j associated with x_j in our random sample will be denoted by $y_{((i))}$ if x_j is corresponding to $x_{(i)}$ in the rearrangement. Consequently we shall have the rearrangement of our sample:

(1.01) $$(x_{(1)}, y_{((1))}),\ (x_{(2)}, y_{((2))}), \ldots\ldots,\ (x_{(n)}, y_{((n))}).$$

According to MAHALANOBIS [1], we may devide the set $\{y_{((j))}\}(j=1,2,\ldots,n)$ into g subsets G_i $(i=1,2,\ldots,g)$ where each G_i consists of $\{y_{((N_{i-1}+k))}\}$ $(k=1,2,\ldots,n'_i)$, such that $N_0=0$, $N_{i-1}=n'_1+n'_2+\ldots+n'_{i-1}$ $(i=1,2,\ldots,g)$ with $n'_1+n'_2+\ldots+n'_g=n$, and let us define

(1.02) $$\bar{y}_{((i))}=(y_{((N_{i-1}+1))}+y_{((N_{i-1}+2))}+\quad +y_{((N_i))})/n'_i.\ (i=1,2,\ldots,g)$$

In what follows simply for the sake of simplicity we are mainly (not exclusively) concerned with the case when $n'_1=n'_2=\ldots=n'_g=n'$ and $gn'=n$.

1.2. Let the probability density function of our bivariate population be denoted by $f(x,y)$, and let us denote the marginal probability distribution density of X by

(1.03) $$f_1(x)=\int_{-\infty}^{\infty} f(x,y)\ dy$$

and the conditional probability density by $f_1(y\,|\,x)$ by

(1.04) $$f_1(y\,|\,x)=f(x,y)/f_1(x).$$

The probability distribution of X is then given by

(1.05) $$F_1(x)=\int_{-\infty}^{x} f_1(x)\ dx=\int_{-\infty}^{x} dx\int_{-\infty}^{\infty} f(x,y)dy.$$

We have consequently

(1.06) $$Pr.\{y<y_{((k))}<y+dy,\ x<x_{(k)}<x+dx\}$$
$$=\frac{n!}{(k-1)!\,(n-k)!}\ (F_1(x))^{k-1}\, f(x,y)\,(1-F_1(x))^{n-k}dxdy$$

(1.07) $$Pr.\{y<y_{((k))}<y+dy\}$$
$$=\frac{n!}{(k-1)!\,(n-k)!}\left(\int_{-\infty}^{\infty}(F_1(x))^{k-1}f(x,y)\,(1-F_1(x))^{n-k}\,dx\right)dy$$

and hence, for $l\geqq 1$,

(1.08) $$E\{y^l_{((k))}\}$$
$$=\int_{-\infty}^{\infty} y^l\ dy\ \frac{n!}{(k-1)!\ (n-k)!}\int_{-\infty}^{\infty}(F_1(x))^{k-1}f(x,y)\,(1-F_1(x))^{n-k}dx,$$

provided that the mements are assumed to exist.

1.3. The simultaneous probability density density function of two statistics $y_{((i))}$ and $y_{((j))}$ is now given by

(1.09) $$g(y_i, y_j)dy_idy_j$$
$$\equiv Pr.\{y_i<y_{((i))}<y_i+dy_i,\ y_j<y_{((j))}<y_j+dy_j\}$$

$$= \frac{n!}{(i-1)!(j-i-1)!(n-j)!} \int_{-\infty}^{\infty} dx_1 \int_{x_1}^{\infty} (F(x_1))^{i-1} (F(x_2)-F(x_1))^{j-i-1}$$

$$(1-F(x_2))^{n-j} f(x_1, y_i) f(x_2, y_j)\, dx$$

for $i<j$.

A product moment of $y_{(i)}$ and $y_{(j)}$ is now given by

(1.10) $$E\left\{y_{(i)}^p\, y_{(j)}^q\right\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y_i^p y_j^q g(y_i, y_j)\, dy_i dy_j,$$

provided that it is assumed to exist.

1.4. Now let us consider two random samples of the same size n from a bivariate population, which we denote by $\{(x_i, y_i)\}$ $(i=1,2,\dots,n)$ and $\{(x_i', z_i)\}$ $(i=1,2,\dots,n)$ respectively. According to the statistical procedure defined in § 1.1, we can define $(\bar{y}_{(1)}, \bar{y}_{(2)}, \dots, \bar{y}_{(g)})$ and $(\bar{z}_{(1)}, \bar{z}_{(2)}, \dots, \bar{z}_{(g)})$.

We are interested with the area S defined in the following way according to MAHALANOBIS [1].

First of all let the k th subarea S_k be defined with respect to the four points:

(1.11) $$(kd, \bar{y}_{(k)}),\ ((k+1)d, \bar{y}_{(k+1)}),\ (kd, \bar{z}_{(k)}),\ ((k+1)d, \bar{z}_{(k+1)}),$$

and then let us define S by

(1.12) $$S = S_1 + S_2 + \dots + S_{g-1}.$$

Now let each S_k be defined in the following way. There are two cases to be distinguished with each other. Let us consider two straight lines

(1.131) $$y - \bar{y}_{(k)} = \frac{\bar{y}_{(k+1)} - \bar{y}_{(k)}}{d}\,(x-kd)$$

(1.132) $$y - \bar{z}_{(k)} = \frac{\bar{z}_{(k+1)} - \bar{z}_{(k)}}{d}\,(x-kd).$$

Let the abscissa of the point where the two straight lines meet with each other be denoted by $x_{k,k+1}$. For our purpose the case (1°) : $x_{k,k+1}<kd$ or $x_{k,k+1}>(k+1)d$, and the case (2°) : $kd<x_{k,k+1}<(k+1)d$, must be distinguished with each other.

The case (1°) : This case will happen under the following two mutually exclusive cases

(1.14) $$\begin{cases} A_1: & \bar{y}_{(k)} \geqq \bar{z}_{(k)}\ \ \bar{y}_{(k+1)} \geqq \bar{z}_{(k+1)} \\ A_2: & \bar{y}_{(k)} < \bar{z}_{(k)}\ \ \bar{y}_{(k+1)} < \bar{z}_{(k+1)}\,, \end{cases}$$

and S_k is defined by

(1.15) $$S_k = d(|\bar{y}_{(k)} - \bar{z}_{(k)}| + |\bar{y}_{(k+1)} - \bar{z}_{(k+1)}|)/2$$

The case (2°) : This case will happen under the following two mutually exclusive cases:

$$(1.16)\qquad \begin{cases} B_1: & \bar{y}_{(\!(k)\!)} \geq \bar{z}_{(\!(k)\!)}, \quad \bar{y}_{(\!(k+1)\!)} < \bar{z}_{(\!(k+1)\!)} \\ B_2: & \bar{y}_{(\!(k)\!)} < \bar{z}_{(\!(k)\!)}, \quad \bar{y}_{(\!(k+1)\!)} > \bar{z}_{(\!(k+1)\!)}, \end{cases}$$

and S_k is given by

$$(1.17)\qquad S_k = \frac{d}{2} \cdot \frac{|\bar{y}_{(\!(k)\!)} - \bar{z}_{(\!(k)\!)}|^2 + |\bar{y}_{(\!(k+1)\!)} - \bar{z}_{(\!(k+1)\!)}|^2}{|\bar{y}_{(\!(k)\!)} - \bar{z}_{(\!(k)\!)}| + |\bar{y}_{(\!(k+1)\!)} - \bar{z}_{(\!(k+1)\!)}|}.$$

Let the simultaneous probability density function of $\bar{y}_{(\!(k)\!)}$ and $\bar{y}_{(\!(k+1)\!)}$ be denoted by $p_k(u,v)$, that is,

$$(1.18)\qquad P\{u < \bar{y}_{(\!(k)\!)} < u+du,\ v < \bar{y}_{(\!(k+1)\!)} < v+dv\} \equiv p_k(u,v)\, dudv.$$

Let us define

$$(1.19)\qquad E\{S_k^l\} \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S_k^l p_k(\bar{y}_k, \bar{y}_{k+1}) p_k(\bar{z}_k, \bar{z}_{k+1}) d\bar{y}_k d\bar{z}_k d\bar{y}_{k+1} d\bar{z}_{k+1}$$

$$= \left(\frac{d}{2}\right)^l \int_{\mathfrak{D}}\int_{(A_1)} ((u-U)+(v-V))^l p_k(u,v) p_k(U,V)\, dudvdUdV$$

$$+ \left(\frac{d}{2}\right)^l \int_{\mathfrak{D}}\int_{(A_2)} ((U-u)+(V-v))^l p_k(u,v) P_k(U,V)\, dudvdUdV$$

$$+ \left(\frac{d}{2}\right)^l \int_{\mathfrak{D}}\int_{(B_1)} \left(\frac{(u-U)^2+(v-V)^2}{(u-U)+(V-v)}\right)^l p_k(u,v) p_k(U,V)\, dudvdUdV$$

$$+ \left(\frac{d}{2}\right)^l \int_{\mathfrak{D}}\int_{(B_2)} \left(\frac{(u-U)^2+(v-V)^2}{(U-u)+(v-V)}\right)^l p_k(u.v) p_k(U,V)\, dudvdUdV$$

$$\equiv I^l(A_1) + I_l(A_2) + I_l(B_1) + I_l(B_2),\ say,$$

provided it does exist, where the domains $\mathfrak{D}(A_i)$ and $\mathfrak{D}(B_i)$ $(i=1,2)$ are defined as follows:

$$(1.201)\qquad \mathfrak{D}(A_1):\ u>U,\ v>V$$

$$(1.202)\qquad \mathfrak{D}(A_2):\ U>u,\ V>v$$

$$(1.203)\qquad \mathfrak{D}(B_1):\ u>U,\ V>v$$

$$(1.204)\qquad \mathfrak{D}(B_2):\ U>u,\ v>V.$$

Let us put

$$(1.21)\qquad p_k^{(1,1)}(u,v) \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_k(x+u, y+v) p_k(x,y)\, dxdy.$$

Then we have, in view of (1.19),

$$(1.22)\qquad I_l(A_1) = \left(\frac{d}{2}\right)^l \int_0^{\infty}\int_0^{\infty} (u+v)^l p_k^{(1,1)}(u,v)\ dudv$$

$$(1.23)\qquad I_l(A_2) = \left(\frac{d}{2}\right)^l \int_0^{\infty}\int_0^{\infty} (u+v)^l p_k^{(1,1)}(-u,-v)\ dudv$$

$$(1.24)\qquad I_l(B_1) = \left(\frac{d}{2}\right)^l \int_0^{\infty}\int_0^{\infty} \left(\frac{u^2+v^2}{u+v}\right)^l p_k^{(1,1)}(u,-v)\ dudv$$

(1.25) $$I_l\ (B_2)=\left(\frac{d}{2}\right)^l \int_0^\infty\int_0^\infty \left(\frac{u^2+v^2}{u+v}\right) p_k^{(1,1)}\ (-u,v)\ dudv.$$

1.5. Let us now calculate the probability density funotion of $\bar{y}_{((k))}$. Let us denote by $E^x_{(h,k)}$ and $E^y_{(h\ k)}$ the events defined by

(1.261) $$E^x_{(h,k)}:\ t_1<x_{(h+1)}<t_1+dt_1,\ t_2<x_{(h+2)}<t_2+dt_2,\dots,$$
$$t_{k-h}<x_{(k)}<t_{k-h}+dt_{k-h}$$

(1.262) $$E^y_{(h,k)}:\ y_1<y_{((h+1))}<y_1+dy_1,\ y_2<y_{((h+2))}<y_2+dy_2,\dots,$$
$$y_{k-h}<y_{((k))}<y_{k-h}+dy_{k-h}$$

respectively. The probability density element for the simultaneous occurrence of $E^x_{(n,k)}$ and $E^y_{(n,k)}$ is given by

(1.27) $$Pr.\ \{E^x_{(n,k)} \cap E^y_{(h,k)}\}$$
$$=\frac{n!}{h!\ (n-k)!}\ (F_1(t_1))^h\ (1-F(t_{k-h}))^{n-k}$$
$$\cdot f(t_1,y_1) f(t_2,y_2),\dots,\ f(t_{k-h},y_{k-h})\ \prod_{i=1}^{k-h} dt_i dy_i\,.$$

Consequently we have

(1.28) $$Pr.\{y<y_{((h+1))}+y_{((h+2))}+\ \dots\ +y_{((k))}<y+dy,\ E^{(x)}_{hk}\}$$
$$=\frac{n!}{n!\ (n-k)!}\ (F_1(t_1))^h(1-F_1(t_{k-h}))^{n-k}\prod_{i=h+1}^{k} f_1\ (t_i)\ dt_i$$
$$\int\int_{y<y_1+y_2+\dots+y_{k-h}<y+dy}\dots\int f(y_1\mid t_1)f(y_2\mid t_2)\dots f(y_{k-h}\mid t_{k-h})dy_1dy_2\dots dy_{k-h},$$

where the integral in the righthand side can be written by convolution

(1.29) $$f_{t_1}{}^*\ f_{t_2}{}^*\dots{}^*\ f_{t_{k-h}}\ (y)dy,$$

where $f_{ti}\equiv f(y\mid t_i)$.

Fininally we have

(1.30) $$Pr.\left\{y<\frac{y_{((h+1))}+y_{((h+2))}+\ \dots\ +y_{((k))}}{k-h}<y+dy\right\}$$
$$=\frac{n!(k-h)}{h!(k-h)!(n-k)!}$$
$$\cdot\int\int_{-\infty<t_1<t_2<\dots t_{k-h}<\infty}\dots\int (F_1(t_1))^h\ (1-F_1(t_{k-h}))^{n-k}$$
$$f_{t_1}{}^*\ f_{t_2}{}^*\dots{}^*\ f_{t_{k-h}}\ ((k-h)y)dt_1\dots dt_{k-h}$$

For $h=(i-1)n'$, $k=in'$, (1.30) gives us the probability distribution element $\bar{y}_{((i))}$, that is,

(1.31) $$Pr.\ \{y<\bar{y}_{((k))}<u+du\}\equiv h_k(u)du$$

$$= \frac{n!\, n'du}{((k-1)n')!n'!((g-k)n')!}$$

$$\int\int\underset{-\infty<t_1<t_2<\cdots<t_{n'}<\infty}{..}\int (F_1(t_1))^{(k-1)n'}(1-F_1(t_{n'}))^{(g-k)n'}\prod_{i=1}^{n'} f_1(t_i)$$

$$f_{t_1}{}^* f_{t_2}{}^* \ldots {}^*{}_{tn'}\ (n'u)\ dt_1 \ldots dt_{n'}.$$

Similarly we have

(1.32) $$Pr.\{u<\bar{y}_{(k)}<u+du,\ v<\bar{y}_{(k+1)}<v+du\} \equiv p_k(u,v)dudv$$

$$= \frac{n!\, n'du\, n'dv}{((k-1)n')!\, n'!n'!\, ((g-k-1)n')!}$$

$$\int\int\underset{-\infty<t_1<t_2<\cdots<t_{2n'}<\infty}{..}\int (F_1(t_1))^{(k-1)n'}(1-F_1(t_{2n'}))^{(g-k-1)n'}\prod_{i=1}^{2n'} f_1(t_i)$$

$$f_{t_1}{}^* f_{t_2}{}^* \ldots {}^* f_{tn'}\ (n'u)$$

$$f_{t_{n'+1}}{}^* f_{t_{n'+2}}{}^* \ldots {}^* f_{t_{2n'}}\ (n'v)dt_1dt_2\ dt_{2n'}.$$

1.6. We shall now enter into the discussion of the separation in the sense of MAHALANOBIS [1]. Let us consider two bivariate populations Π_1 and Π_2 from each of which let us draw independently a random sample of the same size n. Let us denote these two samples from Π_1 and Π_2 by $\{(x_i, y_i)\}$ $(i=1,2,\ldots,n)$ and $\{(x'_i, z_i)\}$ $(i=1,2,\ldots,n)$ respectively.

We can define the area S just as in (1.12) for which the expressions (1.15) and (1.17) hold true under their respective conditions. Now let us define

(1.331) $$Pr.\{y<\bar{y}_{(k)}<y+dy\} \equiv h_{1,k}(y)dy$$

(1.332) $$Pr.\{z<\bar{z}_{(k)}<z+dz\} \equiv h_{2,k}(z)dz$$

(1.341) $$Pr.\{u<\bar{y}_{(k)}<u+du,\ v<\bar{y}_{(k+1)}<v+dv\} \equiv p_{1,k}(u,v)dudv$$

(1.342) $$Pr.\{U<\bar{z}_{(k)}<U+dU, V<\bar{z}_{(k+1)}<V+dV\} \equiv p_{2,k}(U,V)dUdV,$$

in order to distinguish the (possible) difference of the two populations Π_1 and Π_2.

Let us put

(1.35) $$p_k^{(12)}(u,v) \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_{1,k}(x+u, y+v)p_{2,k}(x,y)dxdy.$$

Then the argument sililar to that devoloped in § 1.5 gives us

(1.36) $$E\{S_k^l\} = I_l^{(1,2)}(A_1) + I_l^{(1,2)}\ (A_2) + I_l^{(1,2)}(B_1) + I_l^{(1,2)}(B_2),$$

where

(1.37) $$I_l^{(1,2)}(A_1) = \left(\frac{d}{2}\right)^l \int_0^{\infty}\int_0^{\infty} (u+v)^l\ p_k^{(1,2)}\ (u,v)\ dudv$$

(1.38) $$I_l^{(1,2)}(A_2)=\left(\frac{d}{2}\right)^l\int_0^\infty\int_0^\infty(u+v)^l p_k^{(1,2)}(-u,-v)\,dudv$$

(1.39) $$I_l^{(1,2)}(B_1)=\left(\frac{d}{2}\right)^l\int_0^\infty\int_0^\infty\left(\frac{u^2+v^2}{u+v}\right)p_k^{(1,2)}(u,-v)\,dudv$$

(1.40) $$I_l^{(1,2)}(B_2)=\left(\frac{d}{2}\right)^l\int_0^\infty\int_0^\infty\left(\frac{u^2+v^2}{u+v}\right)p_k^{(1,2)}(-u,v)\,dudv,$$

provided that these integrals are assumed to exist.

1.7. Let us consider two populations Π_1 and Π_2 introduced in § 1.6. The purpose of this paragraph is to decompose the first four moments of the statistic S defined in (1.12) into the sums of some fundamental integrals. For the sake of convenience let us introduce the notations:

(1.41) $$E\{S_k^l\}\equiv(k^l)$$

(1.42) $$E\{S_{k_1}^{l_1}S_{k_2}^{l_2}\}\equiv(k_1^{l_1}k_2^{l_2})$$

(1.43) $$E\{S_{k_1}^{l_1}S_{k_2}^{l_2}S_{k_3}^{l_3}\}\equiv(k_1^{l_1}k_2^{l_3}k_3^{l_3})$$

(1.44) $$E\{S_{k_1}^{l_1}S_{k_2}^{l_2}S_{k_3}^{l_3}S_{k_4}^{l_4}\}\equiv(k_1^{l_1}k_2^{l_2}k_3^{l_3}k_4^{l_4}),$$

where $l_i\geqq 1$ and $1\leqq k_1<k_2<k_3<k_4\leqq g$.

In virtue of these notations we have

(1.45) $$E\{S\}=\sum_{k=1}^{g-1}E\{S_k\}\equiv\sum_{k=1}^{g-1}(k)$$

(1.46) $$E\{S^2\}=\sum_{k=1}^{g-1}E\{S_k^2\}+2\sum_{k=1}^{g-2}E\{S_kS_{k+1}\}+2\sum_{k=1}^{g-3}\sum_{j=k+2}^{g-1}E\{S_kS_j\}$$
$$=\sum_{k=1}^{g-1}(k^2)+2\sum_{k=1}^{g-2}(k(k+1))+2\sum_{k=1}^{g-3}\sum_{j=k+2}^{g-1}(kj)$$

(1.47) $$E\{S^3\}=\sum_{k=1}^{g-1}(k^3)+3\sum_{k=1}^{g-2}(k^2(k+1))+3\sum_{k=1}^{g-2}(k(k+1)^2)$$
$$+3\sum_{k=1}^{g-3}\sum_{j=k+2}^{g-1}(k^2j)+3\sum_{k=1}^{g-3}\sum_{j=k+2}^{g-1}(kj^2)$$
$$+6\sum_{k=1}^{g-3}(k(k+1)(k+2))+6\sum_{k=1}^{g-4}\sum_{j=k+3}^{g-1}(k(k+1)j)$$
$$+6\sum_{k=1}^{g-4}\sum_{j=k+2}^{g-2}(kj(j+1))+6\sum_{k=1}^{g-5}\sum_{j=k+2}^{g-3}\sum_{h=j+2}^{g-1}(kjh)$$

(1.48) $$E\{S^4\}=\sum_{k=1}^{g-1}(k^4)+4\sum_{k=1}^{g-2}(k^3(k+1))+4\sum_{k=1}^{g-2}(k(k+1)^3)$$
$$+12\sum_{k=1}^{g-2}(k^2(k+1)^2)+4\sum_{k=1}^{g-3}\sum_{j=k+2}^{g-1}(k^3j)$$

$$+ 4\sum_{k=1}^{g-3}\sum_{j=k+2}^{g-1}(kj^3)+12\sum_{k=1}^{g-3}\sum_{j=k+2}^{g-1}(k^2j^2)$$

$$+12\sum_{k=1}^{g-3}(k^2(k+1)(k+2))+12\sum_{k=1}^{g-3}(k(k+1)^2(k+2))$$

$$+12\sum_{k=1}^{g-3}(k(k+1)(k+2)^2)+12\sum_{k=1}^{g-4}\sum_{j=k+3}^{g-1}(k^2(k+1)j)$$

$$+12\sum_{k=1}^{g-4}\sum_{j=k+3}^{g-1}(k(k+1)^2j)+12\sum_{k=1}^{g-4}\sum_{j=k+3}^{g-1}(k(k+1)j^2)$$

$$+12\sum_{k=1}^{g-4}\sum_{j=k+2}^{g-1}(k^2j(j+1))+12\sum_{k=1}^{g-4}\sum_{j=k+2}^{g-1}(kj^2(j+1))$$

$$+12\sum_{k=1}^{g-4}\sum_{j=k+2}^{g-1}(kj(j+1)^2)+12\sum_{k=1}^{g-5}\sum_{j=k+2}^{g-3}\sum_{h=j+2}^{g-1}(k^2jh)$$

$$+12\sum_{k=1}^{g-5}\sum_{j=k+2}^{g-3}\sum_{h=j+2}^{g-1}(kj^2h)+12\sum_{k=1}^{g-5}\sum_{j=k+2}^{g-3}\sum_{h=j+2}^{g-1}(kjh^2)$$

$$+24\sum_{k=1}^{g-4}(k(k+1)(k+2)(k+3))+24\sum_{k=1}^{g-5}\sum_{j=k+4}^{g-1}(k(k+1)(k+2)j)$$

$$+24\sum_{k=1}^{g-5}\sum_{j=k+3}^{g-2}(k(k+1)j(j+1))+24\sum_{k=1}^{g-6}\sum_{j=k+3}^{g-3}\sum_{h=j+2}^{g-1}(k(k+1)jh)$$

$$+24\sum_{k=1}^{g-5}\sum_{j=k+2}^{g-3}(kj(j+1)(j+2))+24\sum_{k=1}^{g-6}\sum_{j=k+2}^{g-4}\sum_{h=j+2}^{g-1}(kj(j+1)h)$$

$$+24\sum_{k=1}^{g-6}\sum_{j=k+2}^{g-4}\sum_{h=j+2}^{g-2}(kjh(h+1))+24\sum_{k=1}^{g-7}\sum_{j=k+2}^{g-5}\sum_{h=j+2}^{g-3}\sum_{f=h+2}^{g-1}(kjhf).$$

The mean values given in the right-hand side of (1.45)$\sim$(1.48) can be classified into the following types:

(1) $(k_1^{l_1})$ (2°) $(k^{l_1}(k+1)^{l_2})$ (3°) $(k^{l_1}j^{l_2})$ $(j\geqq k+2)$
(4°) $(k^{l_1}(k+1)^{l_2}(k+2)^{l_3})$ (5°) $(k^{l_1}(k+1)^{l_2}j^{l_3})$
(6°) $(k^{l_1}j^{l_2}(j+1)^{l_3})$ (7°) $(k\,j\,h)$
(8°) $(k^{l_1}(k+1)^{l_2}(k+2)^{l_3}(k+3)^{l_4})$ (9°) $(k^{l_1}(k+1)^{l_2}(k+2)^{l_3}j^{l_4})$
(10°) $(k^{l_1}j^{l_2}(j+1)^{l_3}(j+2)^{l_4})$ (11°) $(k^{l_1}(k+1)^{l_2}j^{l_3}(j+1)^{l_4})$
(12°) $(k^{l_1}(k+1)^{l_2}j^{l_3}h^{l_4})$ (13°) $(k^{l_1}j^{l_2}(j+1)^{l_3}h^{l_4})$
(14°) $(k^{l_1}j^{l_2}h^{l_3}(h+1)^{l_4})$ (15°) $(k^{l_1}j^{l_2}h^{l_3}f^{l_4})$,

where we assume $j\geqq k+2$, $h\geqq j+2$ and $f\geqq h+2$.

1.8. In order to evaluate the mean values given in § 1.7, let us introduce various notations which will be convenient in dealing with them. Let us define for each population Π_i $(i=1,2)$ the probability density element

(1.49) $$Pr.\{x_1<\bar{y}_{(\!(k+1)\!)}<dx_1, x_2<\bar{y}_{(\!(k_2)\!)}<x_2+dx_2, \dots, x_h<\bar{y}_{(\!(k_h)\!)}<x_h+dx_h ; \Pi_i\}$$
$$=p^{(i)}\begin{pmatrix} x_1 & x_2 \dots x_h \\ k_1 & k_2 \dots k_h \end{pmatrix} dx_1\, dx_2 \dots dx_h,$$

and then define, for $0<u_1, u_2, \dots, u_h<\infty$,

(1.50) $$p^{(12)}\begin{pmatrix} \delta_1 u_1 & \delta_2 u_2 \dots \delta_h u_h \\ k_1 & k_2 \dots k_h \end{pmatrix}$$
$$\equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} p^{(1)}\begin{pmatrix} x_1+{}_1u_1 & x_2+{}_2u_2 \dots x_h+{}_hu_h \\ k_1 & k_2 \quad \dots \quad k_h \end{pmatrix}$$
$$p^{(2)}\begin{pmatrix} x_1 & x_2 & \dots x_h \\ k_1 & k_2 & \dots k_h \end{pmatrix} dx_1 dx_2 \dots dx_h,$$

where each δ_i can take either 1 or -1.

By the decomposition given in § 1.7, it is now sufficient to evaluate the integrals of the types given in $(1°)\sim(15°)$ in § 1.7.

Let us write for the sake of convenience the summation over all the possible combinations of the values $\delta_1, \delta_2, \dots, \delta_h$ by $(\delta_1\delta_2\dots\delta_h)$ and let us denote $du_1du_2\dots du_h=d^hu$. Then we have

(1°) $$E\{S_k^{l_1}\}=\sum_{(\delta_1\delta_2)}\int_0^\infty\int_0^\infty S_k^{l}\, p^{(12)}\begin{pmatrix} \delta_1u_1 & \delta_2u_2 \\ k & k+1 \end{pmatrix} d^2u$$

(2°) $$E\{S_k^{l_1}\, S_{k+1}^{l_2}\}=\sum_{(\delta_1\delta_2\delta_3)}\int_0^\infty\cdots\int_0^\infty S_k^{l_1}\, S_{k+1}^{l_2}\, p^{(12)}\begin{pmatrix} \delta_1u_1 & \delta_2u_2 & \delta_3u_3 \\ k & k+1 & k+2 \end{pmatrix} d^3u$$

(3°) $$E\{S_k^{l_1}\, S_j^{l_2}\}=\sum_{(\delta_1\cdots\delta_4)}\int_0^\infty\cdots\int_0^\infty S_k^{l_1}S_j^{l_2}p^{(12)}\begin{pmatrix} \delta_1u_1 & \delta_2u_2 & \delta_3u_3 & \delta_4u_4 \\ k & k+1 & j & j+1 \end{pmatrix} d^4u$$

(4°) $$E\{S_k^{l_1}S_{k+1}^{l_2}S_{k+2}^{l_4}\}=\sum_{(\delta_1\cdots\delta_4)}\int_0^\infty\cdots\int_0^\infty S_k^{l_1}S_{k+1}^{l_2}S_{k+2}^{l_3}\, p^{(12)}\begin{pmatrix} \delta_1u_1 & \delta_2u_2 & \delta_3u_3 & \delta_4u_4 \\ k & k+1 & k+2 & k+3 \end{pmatrix} d^4u$$

(5°) $$E\{S_k^{l_1}S_{k+1}^{l_2}S_j^{l_3}\}=\sum_{(\delta_1\cdots\delta_5)}\int_0^\infty\cdots\int_0^\infty S_k^{l_1}S_{k+1}^{l_2}S_j^{l_3}\, p^{(12)}\begin{pmatrix} \delta_1u_1 & \delta_2u_2 & \delta_3u_3 & \delta_4u_4 & \delta_5u_5 \\ k & k+1 & k+2 & j & j+1 \end{pmatrix} d^5u$$

(6°) $$E\{S_k^{l_1}S_j^{l_2}S_{j+1}^{l_3}\}=\sum_{(\delta_1\cdots\delta_5)}\int_0^\infty\cdots\int_0^\infty S_k^{l_1}S_j^{l_2}S_{j+1}^{l_3}\, p^{(12)}\begin{pmatrix} \delta_1u_1 & \delta_2u_2 & \delta_3u_3 & \delta_4u_4 & \delta_5u_5 \\ k & k+1 & j & j+1 & j+2 \end{pmatrix} d^5u$$

(7°) $$E\{S_k^{l_1}S_j^{l_2}S_h^{l_3}\}=\sum_{(\delta_1\cdots\delta_6)}\int_0^\infty\cdots\int_0^\infty S_k^{l_1}S_j^{l_2}S_h^{l_3}\, p^{(12)}\begin{pmatrix} \delta_1u_1 & \delta_2u_2 & \delta_3u_3 & \delta_4u_4 & \delta_5u_5 & \delta_6u_6 \\ k & k+1 & j & j+1 & h & h+1 \end{pmatrix} d^6u$$

(8°) $$E\{S_k^{l_1}S_{k+1}^{l_2}S_{k+2}^{l_3}S_{k+3}^{l_4}\}$$
$$=\sum_{(\delta_1\cdots\delta_5)}\int_0^\infty\int_0^\infty\cdots\int_0^\infty S_k^{l_1}S_{k+1}^{l_2}S_{k+2}^{l_3}S_{k+3}^{l_4}\, p^{(12)}\begin{pmatrix} \delta_1u_1 & \delta_2u_2 & \delta_3u_3 & \delta_4u_4 & \delta_5u_5 \\ k & k+1 & k+2 & k+3 & k+4 \end{pmatrix} d^5u$$

(9°) $E\{S_k^{l_1}S_{k+1}^{l_2}S_{k+2}^{l_3}S_j^{l_4}\}$

$$=\sum_{(\delta_1\cdots\delta_6)}\int_0^\infty\int_0^\infty\cdots\int_0^\infty S_k^{l_1}S_{k+1}^{l_2}S_{k+2}^{l_3}S_j^{l_4}\; p^{(12)}\begin{pmatrix}\delta_1u_1 & \delta_2u_2 & \delta_3u_3 & \delta_4u_4 & \delta_5u_5 & \delta_6u_6\\ k & k+1 & k+2 & k+3 & j & j+1\end{pmatrix}d^6u$$

(10°) $E\{S_k^{l_1}S_j^{l_2}S_{j+1}^{l_3}S_{j+2}^{l_4}\}$

$$=\sum_{(\delta_1\cdots\delta_6)}\int_0^\infty\cdots\int_0^\infty S_k^{l_1}S_j^{l_2}\,S_{j+1}^{l_3}S_{j+2}^{l_4}\; p^{(12)}\begin{pmatrix}\delta_1u_1 & \delta_2u_2 & \delta_3u_3 & \delta_4u_4 & \delta_5u_5 & \delta_6u_6\\ k & k+1 & j & j+1 & j+2 & j+3\end{pmatrix}d^6u$$

(11°) $E\{S_k^{l_1}S_{k+1}^{l_2}S_j^{l_3}S_{j+1}^{l_4}\}$

$$=\sum_{(\delta_1\cdots\delta_6)}\int_0^\infty\cdots\int_0^\infty S_k^{l_1}S_{k+1}^{l_2}S_j^{l_3}\,S_{j+1}^{l_4}\cdot p^{(12)}\begin{pmatrix}\delta_1u_1 & \delta_2u_2 & \delta_3u_3 & \delta_4u_4 & \delta_5u_5 & \delta_6u_6\\ k & k+1 & k+2 & j & j+1 & j+2\end{pmatrix}d^6u$$

(12°) $E\{S_k^{l_1}S_{(k+1)}^{l_2}S_j^{l_3}S_h^{l_4}\}$

$$=\sum_{(\delta_1\cdots\delta_7)}\int_0^\infty\cdots\int_0^\infty S_k^{l_1}S_{k+1}^{l_2}S_j^{l_3}S_h^{l_4}\cdot p^{(12)}\begin{pmatrix}\delta_1u_1 & \delta_2u_2 & \delta_3u_3 & \delta_4u_4 & \delta_5u_5 & \delta_6u_6 & \delta_7u_7\\ k & k+1 & k+2 & j & j+1 & h & h+1\end{pmatrix}d^7u$$

(13°) $E\{S_k^{l_1}S_j^{l_2}S_{(j+1)}^{l_3}S_h^{l_4}\}$

$$=\sum_{(\delta_1\cdots\delta_7)}\int_0^\infty\cdots\int_0^\infty S_k^{l_1}S_j^{l_2}S_{j+1}^{l_3}S_h^{l_4}\cdot p^{(12)}\begin{pmatrix}\delta_1u_1 & \delta_2u_2 & \delta_3u_3 & \delta_4u_4 & \delta_5u_5 & \delta_6u_6 & \delta_7u_7\\ k & k+1 & j & j+1 & j+2 & h & h+1\end{pmatrix}d^7u$$

(14°) $E\{S_k^{l_1}S_j^{l_2}S_h^{l_3}S_{(h+1)}^{l_4}\}$

$$=\sum_{(\delta_1\cdots\delta_7)}\int_0^\infty\cdots\int_0^\infty S_k^{l_1}S_j^{l_2}S_h^{l_3}S_{(h+1)}^{l_4}\cdot p^{(12)}\begin{pmatrix}\delta_1u_1 & \delta_2u_2 & \delta_3u_3 & \delta_4u_4 & \delta_5u_5 & \delta_6u_6 & \delta_7u_7\\ k & k+1 & j & j+1 & h & h+1 & h+2\end{pmatrix}d^7u$$

(15°) $E\{S_k^{l_1}S_j^{l_2}S_h^{l_3}S_f^{l_4}\}$

$$=\sum_{(\delta_1\cdots\delta_8)}\int_0^\infty\int\cdots\int_0^\infty S_k^{l_1}S_j^{l_2}S_h^{l_3}S_f^{l_4}\cdot p^{(12)}\begin{pmatrix}\delta_1u_1 & \delta_2u_2 & \delta_3u_3 & \delta_4u_4 & \delta_5u_5 & \delta_6u_6 & \delta_7u_7 & \delta_8u_8\\ k & k+1 & j & j+1 & h & h+1 & f & f+1\end{pmatrix}d^8u.$$

The comblications happen from the fact that

$$S_k=u_1+u_2 \qquad \text{when} \qquad \delta_1\delta_2>0$$

(1.51)

$$=\frac{u_1^2+u_2^2}{u_1+u_2} \qquad \text{when} \qquad \delta_1\delta_2<0$$

and similarly for S_j, S_h and S_f.

This fact makes it necessary to subdivide each of the classes (1°)$\sim$(15°) into subclasses according to the sign change of the sequences $(\delta_1\delta_2)$, $(\delta_1\delta_2\delta_3)$, $\ldots$, and $(\delta_1\delta_2\ldots\delta_8)$.

For instance we have

(1.521)
$$\int_0^\infty\int_0^\infty\int_0^\infty S_k^{l_1}S_{k+1}^{l_2}\ p^{(12)}\begin{pmatrix}u_1 & u_2 & u_3\\ k & k+1 & k+2\end{pmatrix} d^3u$$

$$=\int_0^\infty\int_0^\infty\int_0^\infty (u_1+u_2)^{l_1}(u_2+u_3)^{l_2}\ p^{(12)}\begin{pmatrix}u_1 & u_2 & u_3\\ k & k+1 & k+2\end{pmatrix} d^3u$$

(1.522)
$$\int_0^\infty\int_0^\infty\int_0^\infty S_k^{l_1}S_{k+1}^{l_2}\ p^{(12)}\begin{pmatrix}u_1 & -u_2 & u_3\\ k & k+1 & k+2\end{pmatrix} d^3u$$

$$=\int_0^\infty\int_0^\infty\int_0^\infty \left(\frac{u_1^2+u_2^2}{u_1+u_2}\right)^{l_1}\left(\frac{u_2^2+u_3^2}{u_2+u_3}\right)^{l_2} p^{(12)}\begin{pmatrix}u_1 & -u_2 & u_3\\ k & k+1 & k+2\end{pmatrix} d^3u$$

(1.523)
$$\int_0^\infty\int_0^\infty\int_0^\infty S_k^{l_1}S_{k+1}^{l_2}\ p^{(12)}\begin{pmatrix}u_1 & -u_2 & -u_3\\ k & k+1 & k+2\end{pmatrix} d^3u$$

$$=\int_0^\infty\int_0^\infty\int_0^\infty \left(\frac{u_1^2+u_2^2}{u_1+u_2}\right)^{l_1}(u_2+u_3)^{l_2}\ p^{(12)}\begin{pmatrix}u_1 & -u_2 & -u_3\\ k & k+1 & k+2\end{pmatrix} d^3u$$

(1.524)
$$\int_0^\infty\int_0^\infty\int_0^\infty\int_0^\infty S_k^{l_1}S_h^{l_2}\ p^{(12)}\begin{pmatrix}u_1 & u_2 & -u_3 & -u_4\\ k & k+1 & h & h+1\end{pmatrix} d^4u$$

$$=\int_0^\infty\int_0^\infty\int_0^\infty\int_0^\infty (u_1+u_2)^{l_1}\left(\frac{u_3^2+u_4^2}{u_3+u_4}\right)^{l_2} p^{(12)}\begin{pmatrix}u_1 & u_2 & -u_3 & -u_4\\ k & k+1 & h & h+1\end{pmatrix} d^4u$$

and similarly for other summands given above.

§ 2. The asymptotic form of the multivariate probability density function of $(\bar{y}_{(\!(k_1)\!)}, \bar{y}_{(\!(k_2)\!)}, \dots, \bar{y}_{(\!(k_h)\!)})$

2.1. In order to derive the asymptotic form of the probability density function $p(u_1, u_2, \dots, u_k)$ and $p^{(12)}(u_1, u_2, \dots, u_k)$ we shall appeal to the two asymptotic properties valid under their respective conditions. For the sake of convenience we use

Definition 2.1. Let $\{X_1, X_2, \dots, X_k\}$ and $\{X_1^{(n)}, X_2^{(n)}, \dots, X_k^{(n)}\}$ $(n=1, 2, \dots)$ be a sequence of k-dimensional multivariate stochastic variables, and let it be assumed that for any assigned set of k real numbers $(u_1, u_2, \cdots, u_k)$

$$(1^\circ)\quad Pr.\ \{X_1^{(n)}\leqq u_1,\ X_2^{(n)}\leqq u_2, \dots, X_k^{(n)}\leqq u_k\}$$

$$\equiv F_n(u_1, u_2, \dots, u_k) = \int_{-\infty}^{u_1}\int_{-\infty}^{u_2}\cdots\int_{-\infty}^{u_k} f_n(t_1, t_2, \dots, t_k)\ dt_1 dt_2 \dots dt_k$$

(2.01)
$$(2^\circ)\quad Pr.\ \{X_1\leqq u_1, X_2\leqq u_2, \dots, X_k\leqq u_k\}$$

$$\equiv F(u_1, u_2, \dots, u_k) = \int_{-\infty}^{u_1}\int_{-\infty}^{u_2}\cdots\int_{-\infty}^{u_k} f(t_1, t_2, \dots, t_n) dt_1 dt_2 \dots dt_n$$

$$(3^\circ)\quad \lim_{n\to\infty} f_n\,(u_1, u_2, \ldots, u_k) = f\,(u_1, u_2, \ldots, u_k).$$

Then the probability density function is said to be asymptotically convergent in probability law to the probability density function $f(u_1, u_2, \ldots, u_k)$ and is denoted by

$$(2.02)\qquad f_n\,(u_1, u_2, \ldots, u_k) \overset{(n')}{\simeq} f(u_1, u_2, \ldots, u_k).$$

2.2. The first asymptotic property is concerned with the asymptotic convergence in law of the probability density function $f_{t_1} * f_{t_2} * \ldots f_{t_{n'}}(n'u)$ to the normal probability density. Indeed we shall make

Assumption I. For each assigned value of $(t_1, t_2, \ldots, t_{n'})$ we have

$$(2.03)\qquad f_{t_1} * f_{t_2} * f_{t_{n'}}(n'u)$$

$$\overset{(n')}{\simeq} \frac{1}{\sqrt{2\pi\sigma_2^2\,(1-\rho^2)n'}} \exp\left\{-\frac{n'(u-(\mu_2+\rho\sigma_1\sigma_2^{-1}\bar{t}))^2}{2\sigma_2^2(1-\rho^2)}\right\},$$

where

$$(2.04)\qquad \bar{t} = (t_1 + t_2 + \ldots + t_{n'})/n'.$$

The practical uses of this assumption are not only concerned with large n', but also with moderate size of n', as may be expected from the central limit theorem.

Example 2.1. Let us consider the bivariate normal population or populations of the type

$$(2.05)\qquad f\,(x, y) = \frac{1}{2\pi\sigma_1\sigma_2(1-\rho^2)^{1/2}} \exp\{-Q\},$$

where

$$(2.06)\qquad Q \equiv \frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x\mu-{}_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right).$$

Then the marginal probability density function of x and the conditional probability density function of y for an assigned x are given by

$$(2.07)\qquad f_1(x) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left\{-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right\}$$

$$(2.08)\qquad f(y|x) = \frac{1}{\sqrt{2\pi\sigma_2^2(1-\rho^2)}} \exp\left\{-\frac{(y-\mu_2-(\rho\,\sigma_2/\sigma_1)(x-\mu_1))^2}{2\sigma_2^2(1-\rho^2)}\right\}.$$

Consequently we have, in view of (1.08),

$$(2.09)\qquad E\{y^l_{((k))}\}$$

$$= \frac{n!}{(k-1)!(n-k)!}\int_{-\infty}^{\infty}(F_1(x))^{k-1} f_1(x)\,(1-F_1(x))^{n-k}\left(\int_{-\infty}^{\infty} y^l f(y|x)\,dy\right)dx$$

$$=\frac{n!}{(k-1)!(n-k)!}\int_{-\infty}^{\infty}(F_1(x))^{k-1}f_1(x)(1-F_1(x))^{n-k}$$

$$\cdot\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)+\sigma_2(1-\rho^2)^{1/2}\tau)^l e^{-\tau^2/2}\,d\tau$$

$$=\frac{n!}{(k-1)!(n-k)!}\int_{-\infty}^{\infty}(F_1(x))^{k-1}f_1(x)(1-F_1(x))^{n-k}$$

$$\cdot\sum_{k=0}^{l}{}_lC_h(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1))^{l-h}\sigma_2^h(1-\rho^2)^{h/2}\,E_1\{\tau^h\},$$

where $E_1\{\tau^h\}$ are equal to zero for odd positive integers h, while for even positive integers h we have

$$E_1\{\tau^h\}=1.3.5\ldots\ldots(h-1), \tag{2.01}$$

and $E_1\{\tau^h\}=1$.

The transformation $t=(x-\mu_1)/\sigma_1$ yields us therefore

$$E\{y_{(k)}^l\}=\frac{n!}{(k-1)!(n-k)!}\int_{-\infty}^{\infty}(\Phi(t))^{k-1}\varphi(t)(1-\Phi(t))^{n-k} \tag{2.11}$$

$$\cdot\left\{\sum_{h=0}^{l}{}_lC_h(\mu_2+\rho\sigma_2 t)^{l-h}\,\sigma_2^h(1-\rho^2)^{h/2}E_1\{\tau^h\}\right\}\,dt,$$

where we have put

$$\varphi(t)=\frac{1}{\sqrt{2\pi}}\,e^{-t^2/2} \tag{2.121}$$

$$\Phi(t)=\int_{-\infty}^{t}\varphi(u)\,du. \tag{2.122}$$

The special case when $\mu_2=0$ may be worth while for us to mention. Indeed we have then

$$E\{y_{(k)}^l\}=\sigma_2^l\rho^l\sum_{h=0}^{l}{}_lC_h\left(\frac{1-\rho^2}{\rho^2}\right)^{h/2}E_1\{\tau^h\}E\{\sigma_{(k)}^{l-h}\}, \tag{2.13}$$

where we have put

$$E\{X_{(k)}^s\}=\frac{n!}{(k-1)!(n-k)!}\int_{-\infty}^{\infty}(\Phi(t))^{k-1}t^s\varphi(t)(1-\Phi(t))^{n-k}dt. \tag{2.14}$$

The similar argument gives us that, for $i<j$, in view of (1.09),

(2.15) $E\{y_{(i)}^l y_{(j)}^m\}$

$$= \frac{n!}{(i-1)!(j-i-1)!(n-j)!} \int_{-\infty}^{\infty} dx_1 \int_{x_1}^{\infty} (F_1(x_1))^{i-1} f_1(x_1) (F_1(x_2)-F_1(x_1))^{j-i-1} f_1(x_2)(1-F_1(x_2))^{n-j} dx_2$$

$$\cdot \int_{-\infty}^{\infty} y_1^l f(y_1|x_1)\, dy_1 \int_{-\infty}^{\infty} y_2^m f(y_2|x_2)\, dy_2$$

$$= \frac{n!}{(i-1)!(j-i-1)!(n-j)!} \int_{-\infty}^{\infty} dx_1 \int_{x_1}^{\infty} (F_1(x_1))^{i-1} f_1(x_1) (F_1(x_2)-F_1(x_1))^{j-i-1} f_1(x_2)(1-F_1(x_2))^{n-j} dx_2$$

$$\cdot \sum_{h=0}^{l} {}_lC_h \left(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x_1-\mu_1)\right)^{l-h} \sigma_2^h (1-\rho^2)^{h/2} E_1\{\tau^h\}$$

$$\cdot \sum_{g=0}^{m} {}_mC_g \left(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x_2-\mu_2)\right)^{m-g} \sigma_2^g (1-\rho^2)^{g/2} E_1\{\tau^g\}.$$

In particular for $\mu_2=0$, we have

(2.16) $E\{y_{((i))}^l y_{((j))}^m\}$

$$= \sigma_2^{l+m}\rho^{l+m} \sum_{h=0}^{l}\sum_{g=0}^{m} {}_lC_h \cdot {}_mC_g \left(\frac{1-\rho^2}{\rho^2}\right)^{\frac{h+g}{2}} E_1\{\tau^h\}E_1\{\tau^g\}E\{X_{(i)}^{l-h} X_{(j)}^{m-g}\},$$

where

(2.17) $E\{X_{(i)}^s X_{(j)}^g\}$

$$= \frac{n!}{(i-1)!(j-i-1)!(n-j)!} \int_{-\infty}^{\infty} (\Phi(t_1))^{i-1} t_1^s \varphi(t_1)\, dt_1$$

$$\cdot \int_{t_1}^{\infty} (\Phi(t_2)-\Phi(t_1))^{j-i-1}\, t_2^g \varphi(t_2)(1-\Phi(t_2))^{n-j} dt_2.$$

Our results (2.10) and (2.16) show us that $E\{y_{((k))}^l\}$ and $E\{y_{((i))}^l y_{((j))}^m\}$ can be written in term of the constants associated with the order statistics defined for the standard normal population which have been fully discussed by various authors.

We have also

(2.18) $f_{t_1} * f_{t_2} * \dots f_{t'n}\ (n'u)$

$$= \frac{1}{\sqrt{2\pi n'\sigma_2^2(1-\rho^2)}} \exp\left\{-\frac{\left(u-\left(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(\bar{t}-\mu_1)\right)\right)^2}{2n'\sigma_2^2(1-\rho^2)}\right\}.$$

Consequently we have

(2.19) $Pr.\{u<y_{((k))}<u+du\} \equiv h_k(u)du$

$$= \frac{n!\,du}{((k-1)n')!n'!((g-k)n')!}$$

$$\cdot \int\int\int\cdots\int_{-\infty<t_1<t_2<\cdots<t_{n'}<\infty} (\Phi(t_1))^{(k-1)n'}(1-\Phi(t_{n'}))^{(g-k)n'}\prod_{i=1}^{n'}\varphi(t_i)$$

$$\frac{n'^{1/2}}{(2\pi\sigma_2^2(1-\rho^2))^{1/2}}\,\varphi\left(\frac{n'^{1/2}(u-(\mu_2+\rho\sigma_2\bar{t}))}{\sigma_2(1-\rho^2)^{1/2}}\right)dt_1\ldots dt_{n'},$$

and similarly

(2.20) $$Pr.\ \{u<\bar{y}_{(\!(k)\!)}<u+du,\ v<\bar{y}_{(\!(k+1)\!)}<v+dv\}\equiv p_k(u,v)\,du\,dv$$

$$= \frac{n!\,du}{((k-1)n')!n'!N'!((g-k-1)n')!}$$

$$\cdot \int\int\cdots\int_{-\infty<t_1<t_2<\cdots<t_{2n'}<\infty} (\Phi(t_1))^{(k-1)n'}(1-\Phi(t_{2n'}))^{(g-k-1)n}\prod_{i=1}^{n'}\varphi(t_i)$$

$$\frac{n'^{1/2}}{(2\pi\sigma_2^2(1-\rho^2))^{1/2}}\,\varphi\left(\frac{n'(u-(\mu_2+\rho\sigma_2\bar{t}_1))}{\sigma_2(1-\rho^2)^{1/2}}\right)dt_1\ldots dt_{n'}$$

$$\frac{n'^{1/2}}{(2\pi\sigma_2^2(1-\rho^2))^{1/2}}\,\varphi\left(\frac{n'(v-(\mu_2+\rho\sigma_2\bar{t}_2))}{\sigma_2(1-\rho^2)^{1/2}}\right)dt_{n+1}\ldots dt_{2n'},$$

where we have put

(2.21) $$\bar{t}_k=(t_1+t_2+\ldots+t_{n'})/n';\quad \bar{t}_{k+1}=(t_{n'+1}+\ldots+t_{2n'})/n'.$$

Turning back to our general situation, let us consider the means of n' successive order statistics

(2.22) $$\bar{x}_{(k)}=(x_{((k-1)n'+1)}+x_{((k-1)n'+2)}+\ldots+x_{(kn')})/n'$$

for $k=1,2,\ldots,g-1$. Let us denote by

(2.23) $$r_{n'}\begin{pmatrix}t_1 & t_2 & \ldots & t_h\\ k_1 & k_2 & \ldots & k_h\end{pmatrix}$$

the simultaneous probability density function of the statistics $(\bar{x}_{(k_1)}, \bar{x}_{(k_2)},\ldots, \bar{x}_{(k_h)})$, that is,

(2.24) $$\begin{aligned}&Pr.\ \{t_1<\bar{x}_{(k_1)}<t_1+dt_1,\ t_2<\bar{x}_{(k_2)}<t_2+dt_2,\ldots,\ t_h<\bar{x}_{(k_{h_2})}<t_h+dt_h\}\\ &\equiv r_{n'}\begin{pmatrix}t_1 t_2 & \ldots & t_h\\ k_1 k_2 & \ldots & k_h\end{pmatrix}dt_1\,dt_2\ldots dt_h\end{aligned}$$

for any set of h real numbers $(t_1,t_2,\ldots,t_h)$, where $1\leq k_1<k_2<\cdots<k_h<g-1$.

Now it is to be noted that under our Assumption I we have

(2.25) $$\begin{aligned}&Pr.\ \{u_1<\bar{y}_{(\!(k_1)\!)}<u_1+du_1, u_2<\bar{y}_{(\!(k_2)\!)}<u_2+du_2,\ldots,u_h<\bar{y}_{(\!(k_h)\!)}<u_h+du_h\}\\ &\overset{(n')}{\cong}\frac{n'^{h/2}}{(2\pi\sigma_2^2(1-\rho^2))^{h/2}}\int\int_{-\infty}^{\infty}\!\!\cdots\int\prod_{i=1}^{h}\varphi\left(\frac{n'^{1/2}(u_i-(\mu_2+\rho\sigma_2 t_i))}{\sigma_2(1-\rho^2)^{1/2}}\right)\end{aligned}$$

$$r_{n'}\begin{pmatrix} t_1\, t_2 \dots t_h \\ k_1\, k_2 \dots k_h \end{pmatrix} dt_1 dt_2 \dots dt_h.$$

2.3. The second asymptotic property is to appeal to the asymptotic normality of the order statistic. In what follows let us make the following

Assumption II. The probability density function $f_1(x)$ is differentiable in the closure of its carrier $c(f_1)$ and $f_1'(x) \neq 0$ in the interior of $c(f_1)$.

Under our Assumption II we can appeal to the theorem due to F. MOSTELLER [1].

Theorem 1. (F. MOSTELLER) Let m be any assinged positive integers and let $\{\lambda_i\}$ $(i=1,2,\dots,m)$ by any assigned set of m fractional numbers $0<\lambda_1<\lambda_2<<\lambda_m<1$. and let the λ_i-quantile of the population be denoted by τ_i, *i.e.*,

$$F_1(\tau_i) \equiv \int_{-\infty}^{\tau_i} f_1(t)dt = \lambda_i \qquad (i=1,2,\dots,m). \tag{2.26}$$

Let $n_i = [n\lambda_i]+1$, $i=1, 2, ,\dots, m$, where $[x]$ denotes the largest integer not greater than x.

Then as n tends infinity the simultaneous probability distribution of the m order statistics $(x_{(n_1)}, x_{(n_2)}, \dots, x_{(n_m)})$ tends to a k-dimensional normal distribution with the means $(\tau_1, \tau_2, \dots, \tau_m)$ and the variance-covariance $m\times m$ matrix $(\lambda_i(1-\lambda_j)/nf_1(\tau_i)\, f(\tau_j))$ $(i,j=1,2,\dots,m)$.

Let us apply this theorem to the approximation of the integral (2.25). Let us consider the set of fractional numbers $\{\lambda_{k,i}^{(n)}\}$, $\{\tau_{k,i}^{(n)}\}$ and $\{g_{k,i}^{(n)}\}$ $(k=1, 2,\dots, g-1;\ i=1,2,\dots,n')$ such that

$$\lambda_{k,i}^{(n)} \equiv ((k-1)n'+i)/n, \tag{2.27}$$

$$F_1(\tau_{k,i}^{(n)}) \equiv \int_{-\infty}^{\tau_{k,i}^{(n)}} f_1(t)dt = \lambda_{k,i}^{(n)} \tag{2.28}$$

$$g_{k,i}^{(n)} \equiv f_1(\tau_{k,i}^{(n)}). \tag{2.29}$$

The direct application of Mosteller's theorem gives us for any assigned h, $0<h<g-1$

Lemma 2.2. As n' tends to infinity the h-dimensional multivariate statistic $(\bar{x}_{(k_1)}, \bar{x}_{(k_2)}, \dots, \bar{x}_{(k_h)})$ is asymptotically distributed in the h-dimentional normal distribution with the h-dimensional mean vector $(m_{(k_1)}, m_{(k_2)}, \dots, m_{(k_h)})$ and the $h\times h$ variance-covariance matrix$\left(\frac{\sigma_{(k_p,k_q)}}{n'}\right)$ $(p,\ q=1,2,\dots,h)$ such that

$$m_{(k_l)} \equiv \frac{\tau_{k_l,1}^{(n)} + \tau_{k_l,2}^{(n)} + \dots + \tau_{k_l,n'}^{(n)}}{n'} \tag{2.30}$$

$$\frac{\sigma_{(k_p,k_q)}}{n'} \equiv \frac{1}{n'^2}\sum_{i=1}^{n'}\sum_{j=1}^{n'} \frac{\lambda_{k_p,i}^{(n)}(1-\lambda_{k_q,j}^{(n)})}{n g_{k_p,i}^{(n)}\, g_{k_q,j}^{(n)}}. \tag{2.31}$$

Let us write for the sake of convenience

(2.321) $$\sigma_{(k_p,k_p)} \equiv \sigma^2_{(k)}$$

(2.322) $$\rho_{(k_p,k_q)} \equiv \sigma_{(k_p,k_q)}/\sigma_{(k_p)}\sigma_{(k_q)}.$$

2.4. The combination of Lemma 2.1 and 2.2 gives us

Theorem 2.1. Under a situation where both Assumptions I and II hold true, the simultaneous probability density function of $(\bar{y}_{\langle\langle k_1\rangle\rangle}, \bar{y}_{\langle\langle k_2\rangle\rangle}, \ldots, \bar{y}_{\langle\langle k_h\rangle\rangle})$ is asymptotically convergent in probability law, as n' tends to infinity, to the h-variate normal distribution with the mean vector $(\xi_{\langle\langle k_1\rangle\rangle}, \xi_{\langle\langle k_2\rangle\rangle}, \ldots, \xi_{\langle\langle k_h\rangle\rangle})$ and the $h\times h$ variance-covariance matrix $(\rho_{\langle\langle k_i,k_j\rangle\rangle}\ \sigma_{\langle\langle k_i\rangle\rangle}\ \sigma_{\langle\langle k_j\rangle\rangle})$ $(i,j=1,2,\ldots,h)$ such that

(2.33) $$\xi_{\langle\langle k_i\rangle\rangle} \equiv \mu_2 - \rho m_{(k_i)}\ \sigma_2$$

(2.34) $$\sigma^2_{\langle\langle k\rangle\rangle} \equiv \sigma_{\langle\langle k,k\rangle\rangle} \equiv \sigma^2_2\left(\frac{1-\rho^2}{n'_k} + \rho^2\ \sigma^2_{(k)}\right)\quad (k=1\sim5)$$

(2.35) $$\sigma_{\langle\langle k_1,k_2\rangle\rangle} \equiv \rho^2\ \sigma_{(k_1,k_2)}\qquad (k_1 \neq k_i\ ;\ k_1,\ k_2=1\sim5)$$

(2.36) $$\rho_{\langle\langle k_1,k_2\rangle\rangle} \equiv \frac{\sigma_{\langle\langle k_1,k_2\rangle\rangle}}{\sigma_{\langle\langle k_1\rangle\rangle}\ \sigma_{\langle\langle k_2\rangle\rangle}}\qquad (k_1,\ k_2=1\sim5),$$

where h is subject to the same condition to Lemma 2.2, $0<h<g-1$.

Proof: In virtue of Assumption I, we have the asymptotic probability density function given by the right-hand side of (2.25). Now let us write

(2,37) $$\bar{y}_{\langle\langle k_i\rangle\rangle} = \mu_2 + \rho\sigma_2\bar{x}_{(k_i)} + n'^{-1/2}\sigma_2(1-\rho^2)^{1/2}\ z_{k_i}$$

$(i=1,2,\ldots,h)$. Then (2.25) gives us that (1°) the h-dimensional variate $(z_{k_1}, z_{k_2}, \ldots, z_{k_n})$ is asymptotically convergent in probability law to the h-dimensional normal variate with the mean vector $(0,0,\ldots,0)$ and the variance-covariance matrix which reduces to the unit-matrix, and that (2°) $(z_{k_1}, z_{k_2}, \ldots, z_{k_h})$ is asymptotically independent to the variate $(\bar{x}_{(k_1)}, \bar{x}_{(k_2)}, \ldots, \bar{x}_{(k_h)})$. Consequently we have, as n' tends to infinity,

(2.38) $$E\{\bar{y}_{\langle\langle k_i\rangle\rangle}\} \cong \mu_2 + \rho\sigma_2 m_{k_i}$$

(2.39) $$\sigma^2\{\bar{y}_{\langle\langle k_i\rangle\rangle}\} \cong \rho^2\ \sigma^2_2\ \sigma^2\ \{\bar{x}_{(k_i)}\} + n'^{-1}\sigma^2_2(1-\rho^2)\sigma^2\{Z_i\}$$

(2.40) $$Cov.\ \{\bar{y}_{\langle\langle k_i\rangle\rangle}, \bar{y}_{\langle\langle k_j\rangle\rangle}\} \cong Cov.\ \{\rho\sigma_2\bar{x}_{(k_i)}, \rho\sigma_2\bar{x}_{(k_j)}\}.$$

Now the application of Lemma 2.2 gives us

(2.41) $$\sigma^2\ \{\bar{y}_{\langle\langle k_i\rangle\rangle}\} \cong \sigma^2_2(\rho^2\sigma^2_{(k_i)} + (1-\rho^2))/n'$$

(2.42) $$Cov.\{\bar{y}_{\langle\langle k_i\rangle\rangle}, \bar{y}_{\langle\langle k_j\rangle\rangle}\} \cong \rho^2\ \sigma^2_2\ \rho_{(k_i,k_j)}\ \sigma_{(k_i)}\ \sigma_{(k_j)}$$

as we were to prove.

Theorem 2.2. Let Π_1 and Π_2 be two bivariate populations from which two random samples $\{(x_i, y_i)\}$ $(i=1.2,\cdots,n)$ and $\{(x'_i, z_i)\}$ $(i=1,2,\ldots,n)$ are drawn independently. Let $\{(x_{(i)}, y_{\langle\langle i\rangle\rangle})\}$ and $\{(x'_{(i)}, z_{\langle\langle i\rangle\rangle})\}$ be defined as in § 1.6, and then let $(\bar{y}_{\langle\langle k_1\rangle\rangle}, \bar{y}_{\langle\langle k_2\rangle\rangle}, \ldots, \bar{y}_{\langle\langle k_h\rangle\rangle})$ and $(\bar{z}_{\langle\langle k_1\rangle\rangle}, \bar{z}_{\langle\langle k_2\rangle\rangle}, \ldots, \bar{z}_{\langle\langle k_h\rangle\rangle})$ be defined as in § 1.6 respectively for each of these two samples. Let the Assumptions

I and II be satisfied with their respective values of μ_2, ρ, σ_1, σ_2, $\rho_{(k_p,k_q)}$, $\sigma_{(k_p)}$ for each of these two populations, which are denoted by $\mu_2^{(i)}$, $\rho^{(i)}$, $\sigma_1^{(i)}$, $m_{k_p}^{(i)}$, $\sigma_2^{(i)}$ $\rho_{(k_p\ k_q)}^{(i)}$, $\sigma_{(k_p)}^{(i)}$, with $i=1$ for Π_1 and with $i=2$ for Π_2.

Then the simultaneous probability density function of $(\bar{y}_{⟪k_1⟫}-\bar{z}_{⟪k_1⟫}, \bar{y}_{⟪k_2⟫}-\bar{z}_{⟪k_2⟫}, \ldots, \bar{y}_{⟪k_h⟫}-\bar{z}_{⟪k_h⟫})$ is asymptotically convergent in probability law, as n' tend tends to infinity, to the k-variate normal distribution with the mean vector $(\delta_{⟪k_1⟫}, \delta_{⟪k_2⟫}, \ldots, \delta_{⟪k_h⟫})$ and the $h\times h$ variance-covariance matrix $(\sigma^{(1.2)}_{⟪k_i,k_j⟫})$ such that

(2.43) $$\delta_{⟪k_i⟫}\equiv(\mu_2^{(1)}-\rho^{(1)}m^{(1)}_{(k_i)}\sigma_2^{(1)})-(\mu_2^{(2)}-\rho^{(2)}m^{(2)}_{(k_i)}\sigma_2^{(2)}$$

(2.44) $$\sigma^{(1,2)}_{⟪k_i,k_j⟫}\equiv\sigma^{(1)2}_{⟪k_i,k_j⟫}+\sigma^{(2)2}_{⟪k_i,k_j⟫},$$

where

(2.45) $$\sigma^{(1)2}_{⟪k_i,k_j⟫}=\rho^{(1)2}\,\rho^{(1)}_{(k_i,k_j)}\,\sigma^{(1)}_{(k_i)}\,\sigma^{(1)}_{(k_j)}$$

(2.46) $$\sigma^{(2)2}_{⟪k_i,k_j⟫}=\rho^{(2)2}\,\rho^{(2)}_{(k_i,k_j)}\sigma^{(2)}_{(k_i)}\,\sigma^{(2)}_{(k_j)}$$

$(i,j=1,2,\ldots,g-1)$.

Proof: Immediate from Theorem 2.1 and (1.50).

§ 3. The evaluation of the fundamental integrals

3.1. In view of Theorem 2.2, and the results given in § 1.8, the evaluations of the integrals are reduced to those associated with the multivariate normal populations under the Assumptions I and II, provided that we are interested at least with a fairly large or moderately large n'. It is the purpose of this paragraph to enunciate a general procedure to evaluate the integrals of the type:

(3.01) $$I\left(\begin{matrix} l_1 l_2\ldots l_{n-1}\\ j_1 j_2\ldots j_{n-1}\end{matrix}\;\middle|\;\xi_1\,\xi_2,\ldots,\xi_n\right)$$
$$\equiv\frac{|A|^{1/2}}{(2\pi)^{n/2}}\int_0^\infty\int_0^\infty\cdots\int_0^\infty\prod_{i=1}^{n-1}(L_{j_i}(u_i,u_{i+1}))^{l_i}\exp\left\{-\frac{Q}{2}\right\}\prod_{i=1}^{n}du_i,$$

where we have put

(3.02) $$L_{j_i}(u_i,u_{i+1})=u_i+u_{i+1}, \qquad \text{when } j_i=0$$
$$=(u_i^2+u_{i+1}^2)/(u_i+u_{i+1}), \qquad \text{when } j_i=1$$

(3.03) $$Q\equiv\sum_{i=1}^{n}\sum_{j=1}^{n}A_{j_i}(u_i-\xi_i)(u_j-\xi_j)$$

(3.04) $$|A|=|A_{ij}|,$$

with a positive definite quadratic form Q, and non-negative integers l_i $(i=0,1,2,\ldots,n)$.

First let us write

(3.05) $$Q=\sum_{i=1}^{n}\sum_{j=1}^{n}A_{ij}u_iu_j-2\sum_{i=1}^{n}\sum_{j=1}^{n}(A_{ij}\xi_j)u_i+\sum_{i=1}^{n}\sum_{j=1}^{n}A_{ij}\ \xi_i\,\xi_j.$$

Then the change of the order of the integration and the infinite summation gives us

$$(3.06)\qquad I\begin{pmatrix} l_1\, l_2 \dots l_{n-1} \\ j_1\, j_2 \dots j_{n-1} \end{pmatrix}\Big|\, \xi_1\, \xi_2 \dots \xi_n \Big)$$

$$= \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}\, \xi_i\, \xi_j\right\}$$

$$\cdot \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\dots\sum_{kn=0}^{\infty}\prod_{i=1}^{n}\frac{\delta_i^{ki}}{\Gamma(k_i+1)}$$

$$\cdot \frac{|A|^{1/2}}{(2\pi)^{n/2}}\int_0^\infty\int_0^\infty\dots\int_0^\infty u_h^{kn}\prod_{i=1}^{n-1} L_{j_i}(u_i,u_{i+1})^{l_i} u_i^{ki} \exp\left\{-\frac{q}{2}\right\} du_1 \dots du_n,$$

where we have put

$$(3.07)\qquad \delta_i \equiv \sum_{j=1}^{n} A_{ij}\, \xi_j \qquad (i=1, 2, \dots, n)$$

$$(3.08)\qquad q \equiv \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}\, u_i\, u_j.$$

Since q is a positive definite quadratic form, the Jacobi transformation reduces q to its normal form. Indeed let us define

$$(3.09)\qquad D_{i,k} = \begin{vmatrix} A_{11} & A_{12} & \cdots\cdots & A_{1,i-1} & A_{1,k} \\ A_{21} & A_{22} & \cdots\cdots & A_{2,i-1} & A_{2,k} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ A_{i1} & A_{i2} & \cdots\cdots & A_{i,i-1} & A_{i,k} \end{vmatrix} \qquad \begin{matrix} (i=1, 2, \quad \dots, n) \\ (k=i, i+1, \dots, n) \end{matrix}$$

and then let us introce the new system of the variables $(z_1, z_2, \dots, z_n)$ such that

$$(3.10)\qquad z_i = \frac{\sqrt{D_{i,i}}}{\sqrt{D_{i-1,\,i-1}}}\left\{u_i + \sum_{k=i+1}^{n}\frac{D_{i,k}}{D_{i,i}}u_k\right\} \quad (i=1, 2, \dots, n).$$

This yields us

$$(3.11)\qquad q = \sum_{i=1}^{n} z_i$$

$$(3.12)\qquad du_1\, du_2 \dots du_n = D_{n,n}^{-1/2}\, dz_1 \dots dz_n = |A|^{-1/2}\, dz_1 \dots dz_n.$$

The inverse transformation of the transformation $u=(u_1, u_2, \dots, u_n)$ into $z=(z_1, z_2, \dots, z_n)$ is readily seen to be of the form

$$(3.13)\qquad u_i = \sum_{k=i}^{n} C_{i,\,k}\, z_k \qquad (i=1, 2, \dots, n),$$

with a certain non-singular triangular matrix (C_{ik}).

Now the polar coordinate $(r, \theta_1, \theta_2, \dots, \theta_{n-1})$ is defined by

$$(3.14)\qquad z_{1i}=r(\prod_{j=1}^{i-1}\sin\theta^{j})\cos\theta_i\equiv r\varphi_i(\theta)\quad(i=1,2,\ldots,n-1),$$

where $0\leqq\theta_i<2\pi$ $(i=1,2,\ldots,n-2)$, $0\leqq\theta_{n-1}<\pi$, $\theta_n\equiv 0$. This yields us

$$(3.15)\qquad L_0(u_i,u_{i+1})=r\sum_{j=i}^{n}C_{i,j}^{(0)}(\prod_{p=1}^{j-1}\sin\theta_p)\cos\theta_j\equiv r\mathfrak{L}_0^{(i)}(\theta)$$

$$(3.16)\qquad L_1(u_i,u_{i+1})$$

$$=r\frac{(\sum_{j=1}^{n}C_{i,j}(\prod_{p=1}^{j-1}\sin\theta_p)\cos\theta_j)^2+(\sum_{j=i+1}^{n}C_{i+1,j}(\prod_{p=1}^{j=1}\sin\theta_p)\cos\theta_j)^2}{\sum_{k=j}^{n}C_{i,j}^{(0)}(\prod_{p=1}^{i-1}\sin\theta_p)\cos\theta_j}$$

$$\equiv r\,\mathfrak{L}_1^{(i)}(\theta),$$

where we have put

$$(3.17)\qquad C_{i,i}^{(0)}=C_{i,i},\quad C_{i,k}^{(0)}=C_{i,k}+C_{i+1,k}\qquad(k\geqq i+1).$$

We have also

$$(3.18)\qquad \prod_{i=1}^{n}u_i^{ki}=r^{k_1+k_2+\cdots+kn}\prod_{i=1}^{n}(\sum_{j=i}^{n}C_{ij}(\prod_{p=1}^{j-1}\sin\theta_p)\cos\theta_j)^{ki}$$

$$\equiv r^{k_1+k_2+\cdots+kn}\,\mathfrak{N}(\theta;k_1,k_2,\ldots,k_n).$$

The use of (3.11)$\sim$(3.18) in the right-hand side of (3.06) yields us, after integration with respect to r over the domain $0\leqq r<\infty$, that

$$(3.19)\qquad I\begin{pmatrix}l_1 & l_2 & \ldots & l_{n-1}\\ j_1 & j_2 & \ldots & j_{n-1}\end{pmatrix}\Bigg|\ \xi_1\ \xi_2\ \ldots\ \xi_n\Bigg)$$

$$=\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}A_{ij}\,\xi_i\,\xi_j\right\}\cdot\frac{2\sum_{i=1}^{n-1}l_i/2-1}{\pi^{n/2}}$$

$$\cdot\sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\cdots\sum_{k_n=0}^{\infty}\left(\prod_{i=1}^{n}\frac{(2^{1/2}\delta_i)^{k_i}}{(k_i+1)}\right)\Gamma\left(\frac{\sum_{i=1}^{n-1}l_i+\sum_{i=1}^{n}k_i+n}{2}\right)$$

$$\cdot\frac{1}{(2\pi)^{n/2}}\iiint_{\mathfrak{D}}\prod_{i=1}^{n-1}(\mathfrak{L}_{ji}^{(i)}(\theta))^{\,li}\,\mathfrak{N}(\theta;k_1,k_2,\ldots,k_n)$$

$$\cdot\prod_{j=1}^{n-2}\sin^{n-1-j}\theta^j\prod_{i=1}^{n-1}d\theta_i,$$

where the domain of integration $\mathfrak{D}$ in $(\theta_1,\theta_2,\ldots,\theta_{n-1})$ is defined as the domain satisfying the follwing inequalities simultaneously:

$$(3.20)\qquad \sum_{k=i}^{n}C_{i,k}\left(\prod_{j=1}^{k-1}\sin\theta_j\right)\cos\theta_k\geqq 0\qquad(i=1,2,\ldots,n).$$

The integrand in the right-hand side of (3.19) is a rational function of $\sin\theta_j$ and $\cos\theta_j$ $(j=1,2,\ldots,n-1)$, and therefore can be evaluated by the combination of elementary integrations, although there remain a lot of complexity in their individual cases.

3.2. Let us apply our procedure to the evaluations of the special cases of (3.01) when $n=2$. For the convenience let us write

(3.21) $$I\left(\begin{smallmatrix}1\\0\end{smallmatrix}\middle|\,\xi_1,\xi_2\right)\equiv M_l\ (\xi_1,\xi_2\,;\ \sigma_1^2,\sigma_2^2,\rho)$$
$$\equiv\frac{1}{2\pi\sigma_1\sigma_2(1-\rho^2)^{1/2}}\int_0^\infty\int_0^\infty (u+v)^l\exp\left\{-\frac{Q}{2}\right\}dudv$$

(3.22) $$I\left(\begin{smallmatrix}1\\0\end{smallmatrix}\middle|\,\xi_1,\xi_2\right)\equiv N_l\ (\xi_1,\xi_2\,;\ \sigma_1^2,\sigma_2^2,\rho)$$
$$\equiv\frac{1}{2\pi\sigma_1\sigma_2(1-\rho^2)^{1/2}}\int_0^\infty\int_0^\infty\left(\frac{u^2+v^2}{u+v}\right)^l\exp\left\{-\frac{Q}{2}\right\}dudv,$$

where

(3.23) $$\frac{Q}{2}\equiv\frac{1}{2(1-\rho^2)}\left\{\frac{(u-\xi_1)^2}{\sigma_1^2}-2\rho\frac{(u-\xi_1)(v-\xi_2)}{\sigma_1\sigma_2}+\frac{(v-\xi_2)^2}{\sigma_2^2}\right\}$$
$$=\frac{1}{2(1-\rho^2)}\left(\frac{\xi_1^2}{\sigma_1^2}-2\rho\frac{\xi_1\xi_2}{\sigma_1\sigma_2}+\frac{\xi_2^2}{\sigma_2^2}\right)$$
$$-\frac{u}{(1-\rho^2)\sigma_1}\left(\frac{\xi_1}{\sigma_1}-\rho\frac{\xi_2}{\sigma_2}\right)-\frac{v}{(1-\rho^2)\sigma_2}\left(\frac{\xi_2}{\sigma_2}-\rho\frac{\xi_1}{\sigma_1}\right)$$
$$+\frac{1}{2(1-\rho^2)}\left(\frac{u^2}{\sigma_1^2}-2\rho\frac{uv}{\sigma_1\sigma_2}+\frac{v^2}{\sigma_2^2}\right).$$

Our method is based upon the expansions to the effect that

(3.24) $$M_l\ (\xi_1,\xi_2\,;\ \sigma_1^2,\sigma_2^2,\rho)$$
$$=\exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{\xi_1^2}{\sigma_1^2}-2\rho\frac{\xi_1\xi_2}{\sigma_1\sigma_2}+\frac{\xi_2^2}{\sigma_2^2}\right)\right\}$$
$$\cdot\sum_{k_1=0}^\infty\sum_{k_2=0}^\infty\frac{\left(\frac{\xi_1}{\sigma_1}-\rho\frac{\xi_2}{\sigma_2}\right)^{k_1}\left(\frac{\xi_2}{\sigma_2}-\rho\frac{\xi_1}{\sigma_1}\right)^{k_2}}{\Gamma(k_1+1)\Gamma(k_2+1)(1-\rho^2)^{k_1+k_2}}$$
$$\cdot\frac{1}{2\pi(1-\rho^2)^{1/2}}\int_0^\infty\int_0^\infty(\sigma_1u+\sigma_2v)^l u^{k_1}v^{k_2}\exp\left\{-\frac{q}{2}\right\}dudv,$$

and

(3.25) $$N_l\ (\xi_1,\xi_2\,;\ \sigma_1^2,\sigma_2^2,\rho)$$
$$=\exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{\xi_1^2}{\sigma_1^2}-2\rho\frac{\xi_1\xi_2}{\sigma_1\sigma_2}+\frac{\xi_2^2}{\sigma_2^2}\right)\right\}$$
$$\cdot\sum_{k_1=0}^\infty\sum_{k_2=0}^\infty\frac{\left(\frac{\xi_1}{\sigma_1}-\rho\frac{\xi_2}{\sigma_2}\right)^{k_1}\left(\frac{\xi_2}{\sigma_2}-\rho\frac{\xi_1}{\sigma_1}\right)^{k_2}}{\Gamma(k_1+1)\Gamma(k_2+1)(1-\rho^2)^{k_1+k_2}}$$

$$\cdot \frac{1}{2\pi(1-\rho^2)^{1/2}} \int_0^\infty \int_0^\infty \left(\frac{\sigma_1^2 u^2+\sigma_2^2 v^2}{\sigma_1 u+\sigma_2 v}\right)^l u^{k_1} v^{k_2} \exp\left\{-\frac{q}{2}\right\} dudv,$$

where we have put

$$\frac{q}{2} \equiv \frac{1}{2(1-\rho^2)}(u^2-2\rho uv+v^2). \tag{3.26}$$

But we have

$$\frac{1}{2\pi(1-\rho^2)^{1/2}} \int_0^\infty \int_0^\infty (\sigma_1 u+\sigma_2 v)^l u^{k_1} v^{k_2} \exp\left\{-\frac{q}{2}\right\} dudv \tag{3.27}$$

$$=\sum_{h=0}^{l} {}_lC_h\, \sigma_1^h\, \sigma_2^{l-h}$$

$$\cdot \frac{1}{2\pi(1-\rho^2)^{1/2}} \int_0^\infty \int_0^\infty u^{k_1+h}\, v^{k_2+l-h} \exp\left\{-\frac{q}{2}\right\} dudv$$

and

$$\frac{1}{2\pi(1-\rho^2)} \int_0^\infty \int_0^\infty \left(\frac{\sigma_1^2 u^2+\sigma_2^2 v^2}{\sigma_1 u+\sigma_2 v}\right)^l u^{k_1} v^{k_2} \exp\left\{-\frac{q}{2}\right\} dudv$$

$$=\sum_{h=0}^{l} {}_lC_h\, \sigma_1^{2h}\, \sigma_2^{2(l-h)} \tag{3.28}$$

$$\cdot \frac{1}{2\pi(1-\rho^2)^{1/2}} \int_0^\infty \int_0^\infty \frac{u^{k_1+2h}\, v^{k_2+2(l-h)}}{(\sigma_1 u+\sigma_2 v)^l} \exp\left\{-\frac{q}{2}\right\} dudv.$$

In the consequence we have to evaluate the integrals in the right hand sides of (3.27). For this purpose let us make use of the polar coordinate (r, θ) defined by

$$r\cos\theta = \frac{1}{\sqrt{2(1+\rho)}}(u+v), \quad r\sin\theta = \frac{1}{\sqrt{2(1-\rho)}}(u-v). \tag{3.29}$$

This transformation yields us

$$\frac{1}{2\pi(1-\rho^2)^{1/2}} \int_0^\infty \int_0^\infty u^{k_1+h}\, v^{k_2+l-h} \exp\left\{-\frac{1}{2(1-\rho^2)}(u^2-2\rho uv+v^2)\right\} dudv \tag{3.30}$$

$$= \frac{\Gamma\left(\dfrac{k_1+k_2+l+2}{2}\right)}{\pi} \sum_{m=0}^{k_1+h} \sum_{n=0}^{k_2+l-h} (-1)^{k_2+l-h-n}\, {}_{k_1+h}C_m \cdot {}_{k_2+l-h}C_n$$

$$\cdot\ (1+\rho)^{m+n/2}\, (1-\rho)^{(k_1+k_2+l-(m+n))/2}\, I_\omega(m+n, k_1+k_2+l-m-n),$$

where $I_\omega(p, q)$ and ω means

(3.31) $$I_\omega(p,q) \equiv \int_0^\omega \cos^p\theta \sin^q\theta \, d\theta$$

(3.32) $$\omega = \sin^{-1}\sqrt{\frac{1+\rho}{2}}.$$

Now concerning the right-hand side of (3.28), we should rather prefer to take ρ instead of ρ in view of (1.39) and (1.40) and let us make use of the polar coordinate (r, θ) defined by

(3.33) $$r\cos\theta = \frac{1}{\sqrt{2(1+\rho)}}(u-v),\ r\sin\theta = \frac{1}{\sqrt{2(1-\rho)}}(u+v)$$

and the auxiliary angle φ_1 defined by

(3.34) $$\sin\varphi_1 = \frac{(2(1+\rho))^{1/2}(\sigma_1-\sigma_2)}{2(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)^{1/2}},\ 0 \leqq \varphi_1 \leqq \frac{\pi}{2}.$$

This device of change of variables gives us now

(3.35) $$\frac{1}{2\pi(1-\rho^2)^{1/2}}\int_0^\infty\int_0^\infty \frac{u^{k_1+2h}\, v^{k_2+2(l-h)}}{(\sigma_1 u+\sigma_2 v)^l} \exp\left\{-\frac{u^2+v^2+2\rho uv}{2(1-\rho^2)}\right\} du dv$$

$$= \frac{\Gamma\left(\frac{k_1+k_2+l+2}{2}\right)}{2^{l/2}\pi(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)^{l/2}}$$

$$\cdot \sum_{m=0}^{k_1+2h}\sum_{n=0}^{k_2+2(l-h)} {}_{k_1+2h}C_m \cdot {}_{k_2+2(l-h)}C_n (-1)^{k_2-n}$$

$$(1+\rho)^{m+n/2}(1-\rho)^{(k_1+k_2+2l-(m+n))/2} \cdot J_\omega(m+n, k_1+k_2+2l-m-n; l, \varphi_1),$$

where $J(p,q;\, l,\varphi)$ means

(3.36) $$J_\omega(p,q;\, l,\varphi) \equiv \int_\omega^{\pi/2} \frac{\cos^p\theta \sin^q\theta}{\sin^l(\theta+\varphi)}\, d\theta.$$

The integrals (3.31) and (3.36) can be evaluated by elementary procedures.

The special case when $\xi_1=\xi_2=0$ which corresponds to the null hypothesis that the two populations have the same population means is worth while to be mentioned here.

In view of (3.24) and (3.30) we have

(3.37) $$M_l(0,0;\, \sigma_1^2, \sigma_2^2, \rho)$$

$$= \frac{\Gamma\left(\frac{l+2}{2}\right)}{\pi} \sum_{h=0}^{l} {}_lC_h\, \sigma_1^h\, \sigma_2^{l-h} \sum_{m=0}^{h}\sum_{n=0}^{l-h} (-1)^{l-h-n}\, {}_hC_m\, {}_{l-h}C_n$$

$$(1+\rho)^{m/2}\ (1-\rho)^{(l-(m+n))/2}\quad I_\omega(m+n,\ l-(m+n)).$$

In particular we have

$$\text{(3.38)}\qquad M_0(0,0\,;\ \sigma_1^2, \sigma_2^2, \rho) = \frac{1}{\pi}\ \sin^{-1}\sqrt{\frac{1+\rho}{2}}$$

$$\text{(3.39)}\qquad M_1(0,0\,;\ \sigma_1^2, \sigma_2^2, \rho) = \frac{1+\rho}{(2\pi)^{1/2}}\cdot\frac{\sigma_1+\sigma_2}{2} + \frac{(1-\rho)^{1/2}(2^{1/2}-(1-\rho)^{1/2})}{(2\pi)^{1/2}}(\sigma_1-\sigma_2)$$

$$\text{(3.40)}\qquad M_2(0,0\,;\ \sigma_1^2, \sigma_2^2, \rho) = \frac{1}{\pi}\Big\{(\sigma_1+\sigma_2)^2(1+\rho)I_\omega(2,0)$$

$$+2(\sigma_1^2-\sigma_2^2)(1-\rho^2)^{1/2}\ I_\omega(1,1)$$

$$+(\sigma_1-\sigma_2)^2\ (1-\rho)\ I_\omega(0,2)\Big\},$$

where

$$\text{(3.411)}\qquad I_\omega(2,0) = \frac{1}{2}\left(\sin^{-1}\sqrt{\frac{1+\rho}{2}} + \frac{1}{2}(1-\rho^2)^{1/2}\right)$$

$$\text{(3.412)}\qquad I_\omega(1,1) = \frac{1+\rho}{2}$$

$$\text{(3.413)}\qquad I_\alpha(0,2) = \frac{1}{2}\ (\sin^{-1}\sqrt{\frac{1+\rho}{2}} - \frac{1}{2}\ (1-\rho^2)^{1/2}).$$

Similarly, in view of (3.25) and (3.35), we have

$$\text{(3.42)}\qquad N_l(0,0\,;\ \sigma_1^2, \sigma_2^2, -\rho)$$

$$= \frac{\Gamma\left(\frac{l+2}{2}\right)}{2^{l/2}\pi(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)^{l/2}}\sum_{h=0}^{l}\ {}_lC_h\ \sigma_1^{2h}\ \sigma_2^{2(l-h)}$$

$$\cdot\sum_{m=0}^{2h}\ \sum_{n=0}^{2(l-h)}\ {}_{2h}C_m\ {}_{2(l-h)}C_n\ (-1)^n(1+\rho)^{m/2}\ (1-\rho)^{(2l-(m+n))/2}$$

$$\cdot\ J_\omega(m, 2l-(m+n)\ ;\ l, \varphi).$$

In particular we have

$$\text{(3.43)}\qquad N(0,0\,;\ \sigma_1^2, \sigma_2^2,\ -\rho) = \frac{1}{\pi}\left(\frac{\pi}{2} - \sin^{-1}\sqrt{\frac{1+\rho}{2}}\right)$$

$$\text{(3.44)}\qquad N(0,0\,;\ \sigma_1^2, \sigma_2^2,\ -\rho)$$

$$= \frac{1}{2^{3/2}\pi^{1/2}(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)^{1/2}}\Big[(\sigma_1^2+\sigma_2^2)\Big\{(1+\rho)\ J_\omega(2,0\,;\ 1,\varphi_1)$$

$$+(1-\rho)J_\omega(0,2\,;\ 1,\varphi_1)$$

$$+2\,(\sigma_1^2-\sigma_2^2)\,(1-\rho^2)^{1/2}J_\omega(1,1\,;\,1,\varphi_1)\Big]$$

(3.45) $N_2(0,0\,;\,\sigma_1^2,\sigma_2^2,-\rho)$

$$=\frac{(1-\rho)^2}{2\pi(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)}\Big[\sigma_1^4\sum_{m=0}^{4}{}_4C_m\Big(\frac{1+\rho}{1-\rho}\Big)^{m/2}J_\omega(m,4-m\,;\,2,\varphi_1)$$

$$+2\,\sigma_1^2\sigma_2^2\sum_{m=0}^{2}\sum_{n=0}^{2}{}_2C_m\cdot{}_2C_n(-1)^n\Big(\frac{1+\rho}{1-\rho}\Big)^{(m+n)/2}J_\omega(m+n,4-m-n\,;\,2,\varphi_1)$$

$$+\sigma_2^4\sum_{n=0}^{4}{}_4C_n\Big(\frac{1+\rho}{1-\rho}\Big)^{n/2}(-1)^n\,J_\omega(n,4-n\,;\,2,\varphi_1)\Big),$$

where $J_\omega(p,q\,;\,l,\varphi_1)$ $(p,q=0,1,2,3,4\,;\,l=1,2,4)$ can be evaluated by elementary calculations.

Specially when $\sigma_1^2=\sigma_2^2$, we have

(3.461) $$M_1(0,0\,;\,\sigma^2,\sigma^2,\rho)=\frac{(1+\rho)\sigma}{(2\pi)^{1/2}}$$

(3.462) $$N_1(0,0\,;\,\sigma^2,\sigma^2,-\rho)=\frac{\sigma}{2\pi^{1/2}}\Big(\frac{1+\rho}{(1-\rho)^{1/2}}\log\frac{\sqrt{2}+\sqrt{1-\rho}}{\sqrt{1+\rho}}-2^{1/2}\rho\Big)$$

(3.463) $$M_2(0,0\,;\,\sigma^2,\sigma^2,\rho)=\frac{2(1+\rho)\sigma^2}{\pi}\Big(\sin^{-1}\sqrt{\frac{1+\rho}{2}}+\frac{(1-\rho^2)^{1/2}}{2}\Big)$$

(3.464) $$N_2(0,0\,;\,\sigma^2,\sigma^2,\rho)=\frac{\sigma^2}{(1-\rho)\pi}((1+\rho+\rho^2)(1-\rho^2)^{1/2}$$

$$-\rho(2+\rho)\Big(\pi-2\sin^{-1}\sqrt{\frac{1+\rho}{2}}\Big)\,.$$

§ 4. Numerical evaluations of variances, covariances, eorrelation-coefficients of the multivariate normal distribution associated with a graphical fractile analysis.

Let us consider the special case when the size of a sample is 19. Let $\lambda=0.05\sim0.950$ [0.05]. To each of these values of λ, let us defined x and z_i by

(4.01) $$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x_i}e^{-t^2/2}dt=\lambda_i\equiv0.05_i$$

(4.02) $$\frac{1}{\sqrt{2\pi}}e^{-x^2_i/2}=z_i\ .$$

The auxiliary table of the 19×19 matrix $(\lambda_i\,(1-\lambda_j)/z_iz_j)$ $(i,j=1,2,\ldots,$

19) is calculated. Our fractile analysis is here concerned with the following three groupings.

Example 4.1. Let us consider the case $n'_1=n'_2=n'_3=n'_4=4$ and $n'_5=3$. For this grouping the modification of (2.27) is involved because the sizes of subgroups are not equal, and we shall define

(4.031) $\lambda^{(19)}_{k,i}=(4(k-1)+i)/n \qquad (k=1\sim4;\ i=1\sim4)$

(4.032) $\lambda^{(19)}_{k,i}=(16+i)/n \qquad (k=5, i=1, 2, 3)$

and the corresponding $x^{(19)}_{k,i}$ and $z^{(19)}_{k,i}$ according to (4.01) and (4.02) respectively. Let us then calculate

$$S_{k_1,k_2}=\sum_{i=1}^{n'k_1}\sum_{j=1}^{n'k_2}\frac{\lambda^{(19)}_{k_1,i}(1-\lambda^{(19)}_{k_2,j})}{z^{(20)}_{k_1,i}\ z^{(20)}_{k_2,j}} \tag{4.04}$$

for $k_1, k_2=1, 2, 3, 4$.

According to (2.31), we calculate

$$\sigma_{(k_1k_2)}=\frac{1}{19}\cdot\frac{S_{k_1,k_2}}{(n'_{k_1}n'_{k_2})^{1/2}}\ (k_1,k_2=1\sim5), \tag{4.05}$$

and hence $\sigma^2_{(k)}=\sigma_{(k,k)}$ and $\rho_{(k_1,k_2)}$ $(k=1, 2, 3, 4, 5)$ according to (2.321)$\sim$(2.322).

Then we proceed to define $\sigma^2_{《k》}$, $\sigma_{《k_1,k_2》}$ and $\rho_{《k_1,k_2》}$ for $|\rho|=0.1\sim0.9$ [0.1] according to (2.34)$\sim$(2.36), where we use now the normalised value of $\sigma^2_i=1$. This example is called as the Case I.

Example 4.2. The Case II treats with the grouping when $n'_1=4$ and $n'_2=3$ $(i=2\sim6)$.

Example 4.3. The Case III treats with the grouping when $n'_1=n'_2=n'_3=5$ and $n'_4=4$.

For each of these two Cases II and III we can calculate $\sigma_{(k_1,k_2)}$, $\rho_{(k_1,k_2)}$, $\sigma^2_{《k》}$, $\sigma^2_{《k_1,k_2》}$ and $\rho_{《k_1,k_2》}$ similarly as in the Case I.

Table 1, (1)$\sim$(3) give the values of $\sigma_{(k_1,k_2)}$ and $\rho_{(k_1,k_2)}$ for the Cases I, II and III respectively Table 2 gives the values of $\sigma_{《k》}$ as functions of $|\rho|$ for these Cases, and Table 3 those of $\rho_{《k_1,k_2》}$.

Table 1, (1). $\sigma_{(k_1,k_2)}$ (upper) and $\rho_{(k_1,k_2)}$ (lower) for the Case I.

	4 (1)	*4* (2)	*4* (3)	*4* (4)	*3* (5)
4(1)	0.44921 1.00000	0.24381 0.64902	0.15284 0.41957	0.10503 0.27318	0.06223 0.14813
4(2)		0.31416 1.00000	0.22998 0.75493	0.15803 0.49154	0.09364 0.26654
4(3)			0.29540 1.00000	0.23218 0.74473	0.13757 0.40383
4(4)				0.32903 1.00000	0.23181 0.64474
3(5)					0.39289 1.00000

Table 1, (2). $\sigma_{(k_1,k_2)}$ (upper) and $\rho_{(k_1,k_2)}$ (lower) for the Case II.

	4 (1)	*3* (2)	*3* (3)	*3* (4)	*3* (5)	*3* (6)
4(1)	0.44921 1.00000	0.22461 0.67121	0.15379 0.47897	0.11425 0.35310	0.08664 0.25156	0.06223 0.14813
3(2)		0.24929 1.00000	0.19099 0.79849	0 14189 0.58865	0.10760 0.41938	0.07729 0.24695
3(3)			0.22949 1.00000	0 18734 0.81004	0.14207 0.57711	0.10204 0.33983
3(4)				0.23306 1.00000	0.19482 0.78530	0.13993 0.46243
3(5)					0.26406 1.00000	0.21540 0.66874
3(6)						0.39289 1.00000

Table 1, (3). $\sigma_{(k_1,k_2)}$ (upper) and $\rho_{(k_1,k_2)}$ (Iower) for the Case III.

	5 (1)	*5* (2)	*5* (3)	*4* (4)
5(1)	0.49877 1.00000	0.26818 0.62854	0.16052 0.41564	0.09082 0.19187
5(2)		0.36500 1.00000	0.26259 0.79484	0.14858 0.36692
5(3)			0.29902 1.00000	0.25752 0.70265
4(4)				0.44921 1.00000

Table 2, (1). $\sigma_{((k))}$ for the Case I.

$\|\rho\|$	$\overset{4}{\sigma}_{((1))}$	$\overset{4}{\sigma}_{((2))}$	$\overset{4}{\sigma}_{((3))}$	$\overset{4}{\sigma}_{((4))}$	$\overset{3}{\sigma}_{((5))}$
0.1	0.50199	0.50064	0.50045	0.50079	0.57787
0.2	0.50791	0.50256	0.50181	0.50315	0.57941
0.3	0.51762	0.50574	0.50407	0.50706	0.58197
0.4	0.53092	0.51016	0.50721	0.51249	0.58554
0.5	0.54754	0.51579	0.51122	0.51938	0.59010
0.6	0.56720	0.52259	0.51608	0.52769	0.59563
0.7	0.58959	0.53051	0.52177	0.53733	0.60209
0.8	0.61441	0.53950	0.52826	0.54825	0.60947
0.9	0.64137	0.54951	0,53551	0.56037	0.61772

Table 2, (2). $\sigma_{((i))}$ for the Case II.

$\|\rho\|$	$\overset{4}{\sigma}_{((1))}$	$\overset{3}{\sigma}_{((2))}$	$\overset{3}{\sigma}_{((3))}$	$\overset{3}{\sigma}_{((4))}$	$\overset{3}{\sigma}_{((5))}$	$\overset{3}{\sigma}_{((6))}$
0.1	0.50199	0.57662	0.57645	0.57648	0.57675	0.57787
0.2	0.50791	0.57443	0.57374	0.57387	0.57495	0.57941
0.3	0.51762	0.57076	0.56920	0.56948	0.57193	0.58197
0.4	0.53092	0 56558	0.56278	0.56328	0.56767	0.58554
0.5	0.54754	0.55886	0.55441	0.55522	0.56215	0.59010
0.6	0.56720	0.55052	0.54401	0.54519	0.55533	0.59563
0.7	0.58959	0.54051	0.53146	0.53310	0.54717	0.60209
0.8	0.61441	0.52872	0.51660	0.51881	0.53759	0.60947
0.9	0.64137	0.51503	0.49922	0.50211	0.52652	0.61772

Table 2, (3). $\sigma_{((k))}$ for the Case III.

$\|\rho\|$	$\overset{5}{\sigma}_{((1))}$	$\overset{5}{\sigma}_{((2))}$	$\overset{5}{\sigma}_{((3))}$	$\overset{4}{\sigma}_{((4))}$
0.1	0.20299	0.20165	0.20099	0.25199
0.2	0.21195	0.20660	0.20396	0.25797
0.3	0.22689	0.21485	0.20891	0.26793
0.4	0.24780	0.22640	0.21584	0.28187
0.5	0.27469	0.24125	0.22476	0.29980
0.6	0.30756	0.25940	0.23565	0.32172
0.7	0.34640	0.28085	0.24852	0.34761
0.8	0.39122	0.30560	0.26338	0.37749
0.9	0.44201	0.33365	0.28021	0.41136

Table 3. $\rho_{((i,j))}=\frac{\sigma_{((i,j))}}{\sigma_{((i))}\sigma_{((j))}}$ $(i\neq j)$ for the Cases I, II and III.

$\lvert\rho\rvert$	(1, 2)			(1, 3)		
	I	II	III	I	II	III
0.1	0.00970	0.00776	0.01326	0.00608	0.00531	0.00795
0.2	0.03821	0.03079	0.05126	0.02399	0.02111	0.03088
0.3	0.08382	0.06842	0.10932	0.05272	0.04698	0.06635
0.4	0.14403	0.11968	0.18116	0.09081	0.08235	0.11105
0.5	0.21583	0.18351	0.26044	0.13650	0.12665	0.16150
0.6	0.29612	0.25895	0.34181	0.18797	0.17942	0.21465
0.7	0.38196	0.34536	0.42131	0.24345	0.24049	0.26807
0.8	0.47075	0.44252	0.49639	0.30138	0.31009	0.32004
0.9	0.56034	0.55077	0.56566	0.36045	0.38904	0.36944

$\lvert\rho\rvert$	(1, 4)			(1, 5)		
	I	II	III	I	II	III
0.1	0.00418	0.00395	0.00402	0.00215	0.00299	—
0.2	0.01644	0.01568	0.01554	0.00846	0.01187	—
0.3	0.03601	0.03488	0.03315	0.01859	0.02634	—
0.4	0.06176	0.06113	0.05498	0.03203	0.04600	—
0.5	0.09233	0.09395	0.07912	0.04815	0.07037	—
0.6	0.12633	0.13301	0.10394	0.06631	0.09902	—
0.7	0.16244	0.17811	0.12825	0.08590	0.13160	—
0.8	0.19955	0.22939	0.15125	0.10636	0.16788	—
0.9	0.23670	0.28736	0.17252	0.12723	0.20782	—

$\lvert\rho\rvert$	(1, 6)			(2, 3)		
	I	II	III	I	II	III
0.1	—	0.00215	—	0.00918	0.00575	0.01304
0.2	—	0.00846	—	0.03648	0.02318	0.05117
0.3	—	0.01859	—	0.08119	0.05291	0.11155
0.4	—	0.03203	—	0.14220	0.09600	0.19006
0.5	—	0.04815	—	0.21804	0.15410	0.28192
0.6	—	0.06631	—	0.30698	0.22957	0.38235
0.7	—	0.08590	—	0.40711	0 32578	0.48703
0.8	—	0.10636	—	0.51645	0.44751	0.59237
0.9	—	0.12723	—	0.63303	0.60167	0.69563

$\lvert\rho\rvert$	(2, 4)			(2, 5)		
	I	II	III	I	II	III
0.1	0.00630	0.00427	0.00659	0.00324	0.00324	—
0.2	0.02500	0.01722	0.02574	0.01286	0.01303	—
0.3	0.05546	0.03929	0.05573	0.02863	0.02967	—
0.4	0.09671	0.07126	0.09410	0.05016	0.05362	—
0.5	0.14748	0.11432	0.13811	0.07691	0.08562	—
0.6	0.20631	0.17018	0.18515	0.10830	0.12670	—
0.7	0.27165	0.24128	0.23300	0.14365	0.17827	—
0.8	0.34195	0.33105	0.27996	0.18227	0.24228	—
0.9	0.41570	0.44442	0.32484	0.22345	0.32140	—

Table 3. $\rho_{((i,j))}=\frac{\sigma_{((i,j))}}{\sigma_{((i))}\sigma_{((j))}}$ $(i\neq j)$ for the Cases I, II and III.

$\|\rho\|$	(2, 6)			(3, 4)		
	I	II	III	I	II	III
0.1	—	0.00232	—	0.00926	0.00564	0.01144
0.2	—	0.00929	—	0.03678	0.02276	0.04491
0.3	—	0.02094	—	0.08175	0.05201	0.09796
0.4	—	0.03734	—	0.14291	0.09455	0.16705
0.5	—	0.05859	—	0.21861	0.15215	0.24802
0.6	—	0.08485	—	0.30692	0.22739	0.33671
0.7	—	0.11637	—	0.40579	0.32399	0.42932
0.8	—	0.15350	—	0.51307	0.44735	0.52270
0.9	—	0.19677	—	0.62670	0.60537	0.61440

$\|\rho\|$	(3, 5)			(3, 6)		
	I	II	III	I	II	III
0.1	0 00476	0.00427	—	—	0.00306	—
0.2	0.01893	0.01723	—	—	0.01228	—
0.3	0.04221	0.03928	—	—	0.02772	—
0.4	0.07412	0.07115	—	—	0.04955	—
0.5	0.11401	0.11396	—	—	0.07798	—
0.6	0.16112	0.16929	—	—	0.11337	—
0.7	0.21458	0.23939	—	—	0.15626	—
0.8	0.27348	0.32739	—	—	0.20742	—
0.9	0.33687	0.43780	—	—	0.26803	—

$\|\rho\|$	(4, 5)			(4, 6)			(5, 6)		
	I	II	III	I	II	III	I	II	III
0.1	0.00801	0.00586	—	—	0.00420	—	—	0.00646	—
0.2	0.03181	0.02362	—	—	0.01683	—	—	0.02586	—
0.3	0.07070	0.05383	—	—	0.03800	—	—	0.05824	—
0.4	0.12360	0.09748	—	—	0.06788	—	—	0.10368	—
0.5	0.18909	0.15604	—	—	0.10677	—	—	0.16233	—
0.6	0.26551	0.23164	—	—	0.15513	—	—	0.23443	—
0.7	0.35110	0.32726	—	—	0.21362	—	—	0.32037	—
0.8	0.44400	0.44704	—	—	0.28323	—	—	0.42075	—
0.9	0.54245	0.59689	—	—	0.36544	—	—	0.53644	—

Literatures

MAHALANOBIS, P. C.: [1] *A method of fractile graphic analysis with some surmises of results,* Lectures in Japan: Fractile Graphic Analysis, Indian Statistical Institute, November 1958, JFG 1~JFG 6.

[2] *National sanple survey: 7th and 9th Rounds,* All-India Rural, Lectures in Japan: Consumption, Indian Statistical Institue, November 1958, JC1–IC7.

MOSTELLER, F.: [1] *On some 'inefficient' statistics,* Ann. Math. Stat., **17** (1941), 377–408.

TAKEUCHI, K.: [1] *On the lecture by Prof. Mahalanobis,* Sugaku no Ayumi, **6** (1958), 196–201 (in Japanese).