ON THE JOINT DISTRIBUTION OF THE MEDIANS IN SAMPLES FROM A MULTIVARIATE POPULATION

By A. M. Mood

University of Texas

It is well known [1] that in the case of a population having a single variate distributed according to a density function satisfying certain general conditions, the median of a sample is asymptotically normally distributed about the population median as a mean. It is the purpose of this paper to extend this result to populations involving more than one variate. Besides the theoretical interest of such a result, there may be some practical value in it when one is dealing with samples from a population for which the median is a more efficient statistic than the mean, as, for example, when the population variance is not finite.

The complexity of the exact distribution of the sample median increases rapidly with the number of variates which describe the population; it is almost impossible to write out completely the distribution for the general case of k variates. For this reason the author has chosen to give first a detailed presentation for the case of two variates, then use a condensed notation to establish the general result. This is a circuitous route, but it seems to be the only feasible one. A condensed notation is necessary for the general case, but presented alone it would be well-nigh incomprehensible.

1. Distribution of the median in two dimensions. An extension of A. T. Craig's [2] geometrical argument will be used to obtain the exact distribution of the sample median. Let us consider two variates x_1 and x_2 with density function $f(x_1, x_2)$ which shall satisfy the following conditions:

1. $f(x_1, x_2) \geq 0$

2. $$\int_{-\infty}^{\infty} f\left(x_1, \frac{1}{N}\right) dx_1 = \int_{-\infty}^{\infty} f(x_1, 0)\, dx_1 + O\left(\frac{1}{N}\right)$$
$$\int_{-\infty}^{\infty} f\left(\frac{1}{N}, x_2\right) dx_2 = \int_{-\infty}^{\infty} f(0, x_2)\, dx_2 + O\left(\frac{1}{N}\right)$$

3. $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_1, x_2)\, dx_1\, dx_2 = 1$$

4. Each of the equations
$$\int_{-\infty}^{\xi}\int_{-\infty}^{\infty} f(x_1, x_2)\, dx_2 dx_1 = \tfrac{1}{2}$$
$$\int_{-\infty}^{\xi}\int_{-\infty}^{\infty} f(x_1, x_2)\, dx_1 dx_2 = \tfrac{1}{2}$$
has a unique real root.

If ξ_1 and ξ_2 are the respective roots of the two equations of this last condition then the point (ξ_1, ξ_2) is defined to be the population median. It will be assumed in what follows that the coordinate system has been so chosen that $\xi_1 = 0 = \xi_2$.

Let a sample of $2n + 1$ elements $(x_{1\alpha}, x_{2\alpha})(\alpha = 1, 2, \cdots, 2n + 1)$ be drawn from this population. The sample median $(\tilde{x}_1, \tilde{x}_2)$ will be defined as an element (not necessarily in the sample) whose x_1 coordinate is the middle, with respect to magnitude, number of the set of numbers $x_{1\alpha}$, and whose x_2 coordinate is the middle number of the set of numbers $x_{2\alpha}$. Now let us compute the probability that the sample median will lie in the rectangle

$$\tilde{x}_i - \tfrac{1}{2}\, d\tilde{x}_i < x_i < \tilde{x}_i + \tfrac{1}{2}\, d\tilde{x}_i \qquad i = 1, 2.$$

This rectangle will be denoted by R''. The remainder of the plane will be divided into eight other regions $R_1, \cdots, R_4'$ as indicated by the dotted lines in Figure 1. The probability that an element will fall in the region $R_i^{(j)}$ will be denoted by

$$p_i^{(j)} = \iint_{R_i^{(j)}} f(x_1, x_2)\, dx_1\, dx_2 .$$

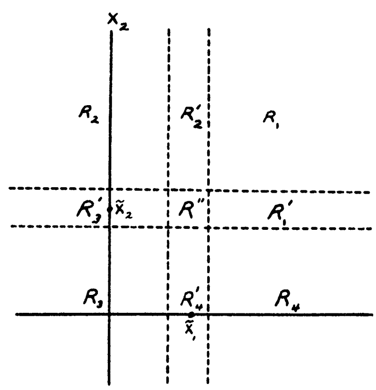

FIG. 1

Neglecting terms involving differentials of higher order we have

$$
\begin{aligned}
p_1 &= \int_{\tilde{x}_1}^{\infty} \int_{\tilde{x}_2}^{\infty} f(x_1, x_2)\, dx_2\, dx_1 \\
p_2 &= \int_{-\infty}^{\tilde{x}_1} \int_{\tilde{x}_2}^{\infty} f(x_1, x_2)\, dx_2\, dx_1 \\
&\vdots \\
p' &= \int_{\tilde{x}_1}^{\infty} f(x_1, \tilde{x}_2)\, dx_1\, d\tilde{x}_2 \\
&\vdots \\
p'' &= f(\tilde{x}_1, \tilde{x}_2)\, d\tilde{x}_1\, d\tilde{x}_2 .
\end{aligned}
\tag{1}
$$

We shall consider now that the sample is drawn from a multinomial population with probabilities $p_1, \cdots, p''$ and pick out those terms which give rise to a sample median in R''. If the median is an element of the sample, then that element must fall in R'' and the other elements must fall in the regions R_1, R_2, R_3, and R_4 in such a manner that

$$n_1 + n_2 = n_3 + n_4 = n$$
$$n_1 + n_4 = n_2 + n_3 = n$$

or so that

$$n_1 = n_3 \text{ and } n_2 = n_4 \tag{2}$$

where n_i is the number of elements in R_i. The probability that this occurs is

$$\sum_{n_1+n_2=n} \frac{(2n+1)!}{n_1!^2\, n_2!^2} p'' p_1^{n_1} p_2^{n_2} p_3^{n_1} p_4^{n_2} \tag{3}$$

Now suppose the median is determined by two different elements of the sample, for example one in R_1' and one in R_2', then there must be n_1 elements in R_1, $n_1 + 1$ elements in R_3, and n_2 elements in each of R_2 and R_4 with

$$n_1 + n_2 = n - 1. \tag{4}$$

The probability in this case is

$$p_1' p_2' \sum_{n_1+n_2=n-1} \frac{(2n+1)!}{n_1!\,(n_1+1)!\,n_2!^2} p_1^{n_1} p_2^{n_2} p_3^{n_1+1} p_4^{n_2}. \tag{5}$$

Continuing in this manner we obtain the distribution of the median, and letting $D(\tilde{x}_1, \tilde{x}_2)$ represent the density function giving this distribution we have

$$\begin{aligned} D(\tilde{x}_1, \tilde{x}_2)\, d\tilde{x}_1\, d\tilde{x}_2 &= p'' \Sigma \frac{(2n+1)!}{n_1!^2\, n_2!^2} (p_1 p_3)^{n_1} (p_2 p_4)^{n_2} \\ &+ (p_3 p_1' p_2' + p_1 p_3' p_4') \Sigma \frac{(2n+1)!}{n_1!\,(n_1+1)!\,n_2!^2} (p_1 p_3)^{n_1} (p_2 p_4)^{n_2} \\ &+ (p_2 p_1' p_4' + p_4 p_2' p_3') \Sigma \frac{(2n+1)!}{n_1!^2\, n_2!\,(n_2+1)!} (p_1 p_3)^{n_1} (p_2 p_4)^{n_2}. \end{aligned} \tag{6}$$

2. Asymptotic distribution of the median in two dimensions. As a simple notation

$$A = B(1 + O(1/\sqrt{n}))$$

will be abbreviated to read

$$A =\cdot B, \tag{7}$$

the dot after the equality sign indicating the omission of the factor $1 + O(1/\sqrt{n})$.

As is customary, the second term of this factor represents any function such that

$$\lim_{N\to\infty} NO(1/N) = L < \infty.$$

In order to get an approximation to (6) for large n we shall use the normal approximation for the multinomial distribution and compute the sums (these cannot be put in finite form) by integration. We use then the well-known result

$$\frac{m!}{\prod_1^r n_i!} \prod_1^r p_i^{m_i} \doteq [A/(2\pi)^{r-1}]^{\frac{1}{2}} \exp\left(-\tfrac{1}{2}\sum_1^{r-1} A_{ij}z_iz_j\right)\prod_1^{r-1} dz_i, \tag{8}$$

where

$$z_i = (m_i - mp_i)/\sqrt{m}, \qquad i = 1, 2, \cdots, r-1, \tag{9}$$

$$A_{ii} = \frac{1}{p_i} + \frac{1}{p_r}, \qquad A_{ij} = \frac{1}{p_r}. \tag{10}$$

Returning to (6) it is to be noted that the fraction immediately following Σ in the first sum has one more factor in the denominator than the corresponding fractions in the other sums. This first sum may therefore be neglected in the asymptotic form as it is of order $1/n$ in comparison with the others. We consider now the second sum in (6) and let it be represented by the letter S

$$S = 2n(2n+1)p_1'p_2' \sum_{n_1+n_2=n-1} \frac{(2n-1)}{n_1!\,(n_1+1)!\,n_2!^2}\, p_1^{n_1}p_2^{n_2}p_3^{n_1+1}p_4^{n_2}. \tag{11}$$

Employing (8) and omitting certain terms of order $1/n$ we have

$$S \doteq 4n^2p_1'p_2' \sum [A/(2\pi)^3]^{\frac{1}{2}} \exp\left(-\tfrac{1}{2}\sum_1^3 A_{ij}z_iz_j\right) dz_1\,dz_2\,dz_3, \tag{12}$$

in which the A_{ij} are defined by (10) with $r = 4$, and

$$z_i = (n_i - 2np_i)/\sqrt{2n}, \qquad i = 1, 2, 3. \tag{13}$$

In view of the relations (2) between the n_i we have

$$\begin{aligned} z_2 &= \sqrt{2n}\,(\tfrac{1}{2} - p_1 - p_2) - z_1 = u_1 - z_1 \\ z_3 &= \sqrt{2n}\,(p_1 - p_3) - z_1 = u_2 - z_1, \end{aligned} \tag{14}$$

in which relations we have defined the new symbols u_1 and u_2. It will be recalled that in (8) the factors dz_i correspond to factors $1/\sqrt{m}$, we therefore let dz_2 and dz_3 in (12) cancel a factor $2n$ from the coefficient of the exponential, and after substituting (14) in (12) find that

$$\begin{aligned} S \doteq{} & 2np_1'p_2' \,\Sigma\, [A/(2\pi)^3]^{\frac{1}{2}} \exp\left\{-\tfrac{1}{2}\left[z_1^2\left(\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} + \frac{1}{p_4}\right)\right.\right. \\ & \left.\left. + 2z_1\left(\frac{u_1+u_2}{p_4} - \frac{u_1}{p_2} + \frac{u_2}{p_3}\right) + \frac{(u_1+u_2)^2}{p_4} + \frac{u_1^2}{p_2} + \frac{u_2^2}{p_3}\right]\right\} dz_1. \end{aligned} \tag{15}$$

The summation can now be performed to within terms of order $1/\sqrt{n}$ by integration with respect to z_1 between the limits $-\infty$ and $+\infty$; this gives us

$$(16)\quad S \doteq \frac{2np_1'p_2'}{2\pi} A^{\frac{1}{2}} \Big/ \left(\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} + \frac{1}{p_4}\right)^{\frac{1}{2}} \exp\left\{-\tfrac{1}{2}\left[\frac{(u_1+u_2)^2}{p_4} + \frac{u_1^2}{p_2} + \frac{u_2^2}{p_3} - \left(\frac{u_1+u_2}{p_4} - \frac{u_1}{p_2} + \frac{u_2}{p_3}\right)^2 \Big/ \left(\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} + \frac{1}{p_4}\right)\right]\right\}.$$

At this point some new symbols are required. We let q_i and q_i' represent the results of replacing $\bar{x}_1$ and $\bar{x}_2$ by zero in the integrals of the relations (1)

$$(17)\quad \begin{aligned} q_1 &= \int_0^\infty \int_0^\infty f(x_1, x_2)\,dx_1\,dx_2 & q_1' &= \int_0^\infty f(x_1, 0)\,dx_1 \\ q_2 &= \int_{-\infty}^0 \int_0^\infty f(x_1, x_2)\,dx_1\,dx_2 & q_2' &= \int_0^\infty f(0, x_2)\,dx_2 \\ q_3 &= \int_{-\infty}^0 \int_{-\infty}^0 f(x_1, x_2)\,dx_1\,dx_2 & q_3' &= \int_{-\infty}^0 f(x_1, 0)\,dx_1 \\ q_4 &= \int_0^\infty \int_{-\infty}^0 f(x_1, x_2)\,dx_1\,dx_2 & q_4' &= \int_{-\infty}^0 f(0, x_2)\,dx_2 \end{aligned}$$

then

$$(18)\quad q_1 + q_2 = q_3 + q_4 = q_1 + q_4 = q_2 + q_3 = \tfrac{1}{2}$$

and

$$(19)\quad q_1 = q_3, \qquad q_2 = q_4.$$

Also we let

$$(20)\quad a_1 = q_2' + q_3', \qquad a_2 = q_1' + q_4',$$

$$(21)\quad y_1 = \sqrt{2n}\, a_1\bar{x}_1, \qquad y_2 = \sqrt{2n}\, a_2\bar{x}_2.$$

We have now

$$(22)\quad \begin{aligned} p_i &\doteq q_i, & i &= 1, 2, 3, 4, \\ p_i' &\doteq q_i' d\bar{x}_2, & i &= 1, 3, \\ p_i' &\doteq q_i' d\bar{x}_1, & i &= 2, 4. \end{aligned}$$

Also

$$\begin{aligned} u_1 &= \sqrt{2n}\,(\tfrac{1}{2} - p_1 - p_2) \\ &= \sqrt{2n} \int_{-\infty}^\infty \int_0^{\bar{x}_2} f(x_1, x_2)\,dx_1\,dx_2 \end{aligned}$$

$$
\begin{aligned}
&= \sqrt{2n}\, \bar{x}_2 \int_{-\infty}^{\infty} f(x_1, \theta\bar{x}_2)\, dx_1, \qquad 0 \le \theta \le 1, \\
&\doteq \sqrt{2n}\, \bar{x}_2 \int_{-\infty}^{\infty} f(x_1, 0)\, dx_1 \\
&\doteq \sqrt{2n}\, a_2 \bar{x}_2 \\
&\doteq y_2 .
\end{aligned} \tag{23}
$$

Similarly

$$
\begin{aligned}
u_2 &= \sqrt{2n}\,(p_1 - p_3) \\
&\doteq -(y_1 + y_2).
\end{aligned} \tag{24}
$$

The result of substituting (22), (23) and (24) in (16) with some further simplification using (18) and (19) is

$$
S \doteq \frac{2nq_1'q_2'}{2\pi\sqrt{q_1q_2}} \exp\left(-\frac{1}{2}\,\frac{y_1^2 - 4(q_1 - q_2)y_1y_2 + y_2^2}{4q_1q_2}\right) d\bar{x}_1\, d\bar{x}_2 . \tag{25}
$$

The other three sums of (6) will give rise to the same expression except that the factors $q_1'q_2'$ will be different; it is clear then that

$$
\begin{aligned}
D(\bar{x}_1, \bar{x}_2)\, d\bar{x}_1\, d\bar{x}_2 &\doteq \frac{2n(q_1'q_2' + q_1'q_2' + q_2'q_3' + q_3'q_4')}{2\pi\sqrt{q_1q_2}} \\
&\qquad \times \exp\left(-\frac{1}{2}\,\frac{y_1^2 - 4(q_1 - q_2)y_1y_2 + y_2^2}{4q_1q_2}\right) d\bar{x}_1\, d\bar{x}_2 \\
&\doteq \frac{2na_1a_2}{2\pi\sqrt{q_1q_2}} \exp\left(-n\,\frac{a_1^2\bar{x}_1^2 - 4(q_1 - q_2)a_1a_2\bar{x}_1\bar{x}_2 + a_2^2\bar{x}_2^2}{4q_1q_2}\right) d\bar{x}_1\, d\bar{x}_2, \qquad (26) \\
&\doteq \frac{1}{2\pi\sqrt{q_1q_2}} \exp\left(-\frac{1}{2}\,\frac{y_1^2 - 4(q_1 - q_2)y_1y_2 + y_2^2}{4q_1q_2}\right) dy_1\, dy_2 . \qquad (27)
\end{aligned}
$$

This is the asymptotic form for the distribution of the median in two dimensions.

3. Distribution of the median in k dimensions. We consider now a population characterized by a density function $f(x_1, \cdots, x_k)$ defined over a euclidean space of k dimensions satisfying conditions like those required of $f(x_1, x_2)$ in section 1, and we assume that the population median is at the origin so that the integral of the density function over any half-space determined by a coordinate hyperplane is $\frac{1}{2}$.

A sample of $2n + 1$ elements will have a median $(\bar{x}_1, \cdots, \bar{x}_k)$ each coordinate of which is the middle number of the set of numbers giving the corresponding coordinate of the elements of the sample. To obtain the probability that the sample median lies in the hyperparallopiped $\bar{x}_\alpha - \frac{1}{2}d\bar{x}_\alpha < x_\alpha < \bar{x}_\alpha + \frac{1}{2}\, d\bar{x}_\alpha$ $(\alpha = 1, 2, \cdots, k)$, we divide the space into 3^k regions by means of hyperplanes perpendicular to the coordinate axes through the points $\tilde{x}_\alpha \pm \frac{1}{2} d\tilde{x}_\alpha$ on the coordinate axes. These regions are illustrated in Figure 2 for the case of three dimensions. The coordinate axes have been omitted in this figure. There will be 2^k primary regions denoted by $R_1, R_2, \cdots, R_{2^k}$ corresponding to the octants of the figure; $k2^{k-1}$ regions with one differential dimension denoted by $R'_1, R'_2, \cdots, R'_{k2^{k-1}}$ corresponding to the quarter slabs of the figure; $\binom{k}{2} 2^{k-1}$ regions with two differential dimensions corresponding to the half strips of the figure, and so forth. Probabilities associated with these regions are defined by

$$p_i^{(j)} = \int_{R_i^{(j)}} f(x_1, \cdots, x_k)\, dx_1 \cdots dx_k .$$

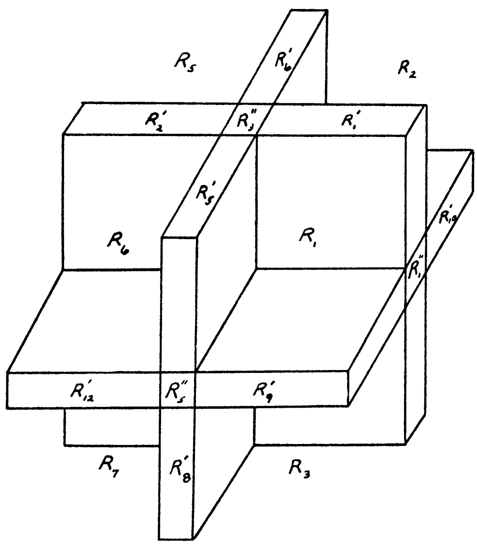

FIG. 2

If the sample median is determined by k different elements of the sample there will be one of these k elements in each of k regions R'_i whose differential dimensions are mutually orthogonal and the other elements of the sample will fall in the regions R_i in such a way that n elements of the sample will lie on either side of any of the k hyperplanes $x_\alpha = \tilde{x}_\alpha$. The probability of this occurrence for a particular choice of k of the regions R'_i is

$$S = \prod_{\alpha=1}^{k} p'_{i_\alpha} \sum \frac{(2n+1)!}{\prod n_i!} \prod_{i=1}^{2^k} p_i^{n_i} \tag{28}$$

in which the 2^k indices n_i are subject to k independent restrictions of the type

$$\Sigma' n_i = n - c_\alpha , \tag{29}$$

where c_α is an integer such that $0 \leq c_\alpha < k$, and the prime on Σ indicates that the sum is to be taken over all n_i on one side of a hyperplane $x_\alpha = \tilde{x}_\alpha$. n_i is the number of elements in R_i and besides the k conditions (29) we have also

$$\sum_1^{2^k} n_i = 2n - k + 1. \tag{30}$$

In order to include all ways in which the median is determined by k different elements of the sample we must add together $2^{k(k-1)}$ sums of the type (28). If the median is determined by less than k elements, say $k - h$ elements, then the fraction $(2n + 1)!/\Pi n_i!$ will have h extra factors in the denominator and hence the sum will be of order $1/n^h$ as compared with that of (28) and may be neglected in obtaining an asymptotic expression.

Thus we need only find the limiting form of (28)

$$S = (2n + 1)(2n) \cdots (2n - k + 2) \prod_1^k p'_{i_\alpha} \sum \frac{(2n - k + 1)!}{\prod n_i!} \prod_1^{2^k} p_i^{n_i},$$

which after substituting (8) and neglecting terms of lower order becomes

$$S \doteq (2n)^k \prod p'_{i_\alpha} \sum (A/(2\pi)^{2^k - 1})^{\frac{1}{2}} \exp\left(-\tfrac{1}{2} \sum A_{ij} z_i z_j\right) \prod_1^{2^k - 1} dz_i, \tag{31}$$

in which the A_{ij} are defined by (10) with $r = 2^k$ and

$$z_i = (n_i - 2np_i)/\sqrt{2n}, \qquad i = 1, 2, \cdots, 2^k - 1. \tag{32}$$

Now we define

$$u_\alpha = \sqrt{2n}(\tfrac{1}{2} - \Sigma' p_i), \qquad \alpha = 1, 2, \cdots, k, \tag{33}$$

the Σ' having the same significance as in (29). These conditions (29) may now be put in the form

$$z_\alpha \doteq u_\alpha - L_\alpha(z),$$

in which $L_\alpha(z)$ is a sum of a certain subset of the variables $z_{k+1}, \cdots, z_{2^k-1}$. Care must be taken in labeling the regions R_i in order to be able to solve for $z_1, \cdots, z_k$ in this form. After substituting these relations in (31) we replace $\prod_1^k dz_\alpha$ by $(1/2n)^{k/2}$ and perform the summation to within terms of order $1/\sqrt{n}$ by integrating the remaining z_i from $-\infty$ to $+\infty$; the result is

$$S \doteq (2n/2\pi)^{k/2} \prod_{\alpha=1}^k p'_{i_\alpha} \sqrt{B} \exp\left(-\tfrac{1}{2} \sum_1^k B_{\alpha\beta} u_\alpha u_\beta\right), \tag{34}$$

in which the $B_{\alpha\beta}$ are functions of the p_i, and $B = | B_{\alpha\beta} |$. As in (17) and (20) we define

$$(35)\qquad \begin{aligned} q_i &= \int_{\bar{R}_i} f(x_1, \cdots, x_k)\Pi\, dx_\alpha \\ q_i' &= \int_{\bar{R}_i'} f(x_1, \cdots, x_k)\Pi'\, dx_\alpha \\ a_\alpha &= \int_{x_\alpha=0} f(x_1, \cdots, x_k)\Pi'\, dx_\alpha = \Sigma' q_i', \end{aligned}$$

in which $\bar{R}_i$ is the set of regions bounded by the coordinate hyperplanes $\bar{R}_i'$ are regions into which the coordinate hyperplanes are divided by the remaining coordinate hyperplanes. Π' indicates that one of the differentials is omitted and the variate corresponding to that differential is put equal to zero in $f(x_1, \cdots, x_k)$; Σ' indicates the sum over all q' determined by regions lying in the hyperplane $x_\alpha = 0$. It is clear that

$$(36)\qquad \begin{aligned} p_i &\doteq q_i \\ \prod_\alpha p_{i_\alpha}' &\doteq \prod_\alpha q_{i_\alpha}' d\tilde{x}_\alpha \\ u_\alpha &\doteq \sqrt{2n} \sum_{\beta=1}^{k} \delta_{\alpha\beta} a_\beta \tilde{x}_\beta = \sum \delta_{\alpha\beta} y_\beta, \end{aligned}$$

where

$$\delta_{\alpha\beta} = \pm 1 \text{ or } 0, \qquad \text{and} \qquad \gamma_\beta = \sqrt{2n} a_\beta \tilde{x}_\beta .$$

Making these substitutions in (34) we have

$$(37)\qquad S \doteq (2n/2\pi)^{k/2} \prod_1^k q_{i_\alpha}' \sqrt{C} \exp\left(-n \sum_1^k C_{\alpha\beta} a_\alpha a_\beta \tilde{x}_\alpha \tilde{x}_\beta\right) \prod d\tilde{x}_\alpha,$$

and adding together all possible sums of the type (28) we have the asymptotic form of the distribution of the sample median

$$(38)\qquad D(\tilde{x}_1, \cdots, \tilde{x}_k) \prod d\tilde{x}_\alpha$$

$$\doteq (2n/2\pi)^{k/2} \prod_1^k a_\alpha \sqrt{C} \exp\left(-n \sum_1^k C_{\alpha\beta} a_\alpha a_\beta \tilde{x}_\alpha \tilde{x}_\beta\right) \prod d\tilde{x}_\alpha$$

$$(39)\qquad \doteq (1/2\pi)^{k/2} \sqrt{C} \exp\left(-\tfrac{1}{2} \sum C_{\alpha\beta} y_\alpha y_\beta\right) \prod dy_\alpha,$$

in which the $C_{\alpha\beta}$ are functions of the q_i.

4. The case of three dimensions. The computation of the coefficients $C_{\alpha\beta}$ of (39) requires the evaluation of a determinant of order $2^k - k$ for each one of them. This work was quite laborious even for $k = 3$ and the author made no attempt to find their explicit expression for larger values of k.

If we let a subscript $+$ indicate integration of the density function $f(x_1, x_2, x_3)$ from 0 to ∞, and a subscript—indicate integration from $-\infty$ to 0, as for example,

$$f_{++-} = \int_0^\infty \int_0^\infty \int_{-\infty}^0 f(x_1, x_2, x_3)\, dx_3\, dx_2\, dx_1,$$

then the q_i of (35) will be defined as follows

$$\begin{aligned} q_1 &= f_{+++} & q_5 &= f_{-++} \\ q_2 &= f_{++-} & q_6 &= f_{-+-} \\ q_3 &= f_{+-+} & q_7 &= f_{--+} \\ q_4 &= f_{+--} & q_8 &= f_{---} \end{aligned} \tag{40}$$

The coefficients $C_{\alpha\beta}$ may be written

$$\begin{aligned} DC_{11} &= 2(q_1 + q_5)(q_2 + q_6) \\ DC_{22} &= 2(q_1 + q_3)(q_2 + q_4) \\ DC_{33} &= 2(q_1 + q_2)(q_3 + q_4) \\ DC_{12} &= q_3q_5 + q_4q_6 - q_1q_7 - q_2q_8 \\ DC_{13} &= q_2q_5 + q_4q_7 - q_1q_6 - q_3q_8 \\ DC_{23} &= q_2q_3 + q_6q_7 - q_1q_4 - q_5q_8, \end{aligned} \tag{41}$$

where

$$\begin{aligned} D = q_1q_2q_3q_4\left(\frac{1}{q_1} + \frac{1}{q_2} + \frac{1}{q_3} + \frac{1}{q_4}\right) &+ q_5q_6q_7q_8\left(\frac{1}{q_5} + \frac{1}{q_6} + \frac{1}{q_7} + \frac{1}{q_8}\right) \\ &+ 2(q_5 + q_6)(q_7 + q_8)(q_1q_2 + q_3q_4) \\ &+ 2(q_5 + q_7)(q_6 + q_8)(q_1q_3 + q_2q_4) \\ &+ 2(q_5 + q_8)(q_6 + q_7)(q_1q_4 + q_2q_3) \\ &+ 8(q_1q_4q_6q_7 + q_2q_3q_5q_8) \end{aligned} \tag{42}$$

(41) and (42) can of course be put in different forms by using the four relations between the q_i. The a_α of (38) are defined in (35); for $k = 3$ they are

$$\begin{aligned} a_1 &= \int_{-\infty}^\infty \int_{-\infty}^\infty f(0, x_2, x_3)\, dx_2\, dx_3 \\ a_2 &= \int_{-\infty}^\infty \int_{-\infty}^\infty f(x_1, 0, x_3)\, dx_1\, dx_3 \\ a_3 &= \int_{-\infty}^\infty \int_{-\infty}^\infty f(x_1, x_2, 0)\, dx_1\, dx_2. \end{aligned} \tag{43}$$

5. The normal distribution in two dimensions. If the density function of the second section of the paper is normal

$$f(x_1, x_2) = 1/(2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}) \exp\left[-\frac{1}{2(1-p^2)}\left(\frac{x_1^2}{\sigma_1^2} - 2\rho\frac{x_1x_2}{\sigma_1\sigma_2} + \frac{x_2^2}{\sigma_2^2}\right)\right], \tag{44}$$

we find that the parameters of (26) are

$$\begin{aligned} q_1 &= \frac{1}{4} + \frac{1}{2\pi}\sin^{-1}\rho, \qquad & q_2 &= \frac{1}{4} - \frac{1}{2\pi}\sin^{-1}\rho, \\ a_1 &= \frac{1}{\sqrt{2\pi}\,\sigma_1}, & a_2 &= \frac{1}{\sqrt{2\pi}\,\sigma_2}. \end{aligned} \tag{45}$$

These give an interesting result—the correlation coefficient of the asymptotic distribution of the sample medians is

$$\rho_m = \frac{2}{\pi}\sin^{-1}\rho \tag{46}$$

hence

$$|\rho_m| \leq |\rho| \tag{47}$$

the equality sign holding only when $\rho = 0$ or ± 1.

REFERENCES

[1] S. S. Wilks, *Statistical Inference*, Ann Arbor, Edwards Bros., 1937.
[2] A. T. Craig, "On the distributions of certain statistics," *Amer. Jour. Math.*, Vol. 54 (1932), pp. 353–366.