ON THE ASYMPTOTIC DISTRIBUTION OF DIFFERENTIABLE STATISTICAL FUNCTIONS

By R. v. Mises

Harvard University

Table of Contents

PAGE

Introduction 309

Part I. Preliminary Theorems.

1. Asymptotically Equal Distributions 311
2. Special Class of Statistical Functions: Quantics 312
3. Asymptotic Expectation of Excess-Power Products 314
4. Asymptotic Expectation and Variance of Quantics 317
5. Final Statement on the Limit of Expectation of Quantics 320
6. Theorem on Products of n Functions 322

Part II. Differentiable Statistical Functions.

1. Definitions 323
2. Taylor Development 325
3. General Theorem 327
4. Illustrations 329

Part III. Second-Type Asymptotic Distribution.

1. Statement of the Problem 331
2. Characteristic Function 332
3. Asymptotic Value of $Q_n(u)$ 335
4. Asymptotic Value of $P_n(x)$ 338
5. Transition to the Continuous Case 342

References 348

Introduction. If n real variables $x_1, x_2, \cdots, x_n$ are subject to a probability distribution with the element $dV_1(x_1)dV_2(x_2) \cdots dV_n(x_n)$ one can ask for the distribution of any function f of $x_1, x_2, \cdots x_n$. We are primarily interested in *statistical functions*, i.e. in functions that depend on the *repartition*[1] $S_n(x)$ of the n quantities $x_1, x_2, \cdots x_n$ only. The simplest case is that of the *linear statistical functions*

$$f = \int \psi(x)\, dS_n(x) = \frac{1}{n}[\psi(x_1) + \psi(x_2) + \cdots + \psi(x_n)]. \tag{1}$$

The so-called Central Limit Theorem of Probability Calculus states that the distribution of a linear statistical function, if n tends to infinity, approaches more and more the normal (Gauss) distribution if some very general conditions linking $\psi(x)$ and the $V_\nu(x)$ are fulfilled. It has been shown, ten years ago, [2] that the restriction to linear functions here is immaterial. Much more general

[1] The function $S_n(x)$ is called the repartition of the real quantities $x_1, x_2, \cdots, x_n$ if $nS_n(x)$ is the number of those among the $x_1, x_2, \cdots, x_n$ that are smaller than or equal to x.

statistical functions tend towards normalcy with increasing n, for example the variance of mth order

$$f = M_m = \int (x - a)^m \, dS_n(x), \qquad a = \int x \, dS_n(x) \tag{2}$$

and, likewise, such combinations as the Lexis quotient $M_2/a(1 - a/N)$ or Gini's disparity measure $1 - \int (1 - S_n)^2 dx/a$ or, in the multidimensional case, the correlation coefficient, etc. On the other hand, statistical functions are known whose distributions assume, asymptotically, a form different from the Gaussian. One example is Pearson's Chi-square, another the test function ω^2, introduced by H. Cramér [1] and the author [4]:

$$f = \omega^2 = \int g'(x)[S_n(x) - \bar{V}_n(x)]^2 \, dx \tag{3}$$

where $g'(x) > 0$ and

$$\bar{V}_n(x) = \frac{1}{n} [V_1(x) + V_2(x) + \cdots + V_n(x)]. \tag{4}$$

N. V. Smirnoff [7, 8] computed the asymptotic distribution of ω^2 for the case that all $V_\nu(x)$ and, therefore, $\bar{V}_n(x)$ equal one and the same distribution function $V(x)$. The result differs widely from the Gaussian distribution.

In order to understand all this it is necessary to consider f as a function defined *in the space of distributions* $V(x)$ (or in a sub-space of it). Then, the variable f whose distribution is sought is the value of $f\{V(x)\}$ at the "point" $S_n(x)$ and should be written as $f\{S_n(x)\}$. Such "functions of functions" were first introduced by Vito Volterra (1887) and are today a familiar topic of higher analysis. The first statement that can be made is that the asymptotic distribution of $f\{S_n(x)\}$ depends mainly on the behavior of $f\{V(x)\}$ at the point $\bar{V}_n(x)$ defined by (4).

Volterra also introduced the notion of derivatives and of Taylor development for a "fonction de ligne." Using these concepts a more specific statement can be pronounced: *The type of asymptotic distribution of a differentiable statistical function* $f\{S_n(x)\}$ *depends on which is the first non-vanishing term in the Taylor development of* $f\{V(x)\}$ *at the point* $\bar{V}_n(x)$; *if it is the linear term the limiting distribution is normal, under restrictions that can easily be derived from the Central Limit Theorem*; *in other cases higher types of asymptotic distributions result.*

The present paper tries to establish this theorem and to furnish preliminary information about the asymptotic distribution of the *second type.*

If both the function $f\{V(x)\}$ and the sequence of distributions $V_1(x)$, $V_2(x)$, $V_3(x)$, $\cdots$ are defined independently of each other, it cannot be presumed that the derivative of f vanishes at $\bar{V}_n(x)$. In this sense the normal distribution appears as the "general case" of an asymptotic distribution while the higher types represent certain "singularities." In the case of type m, ($m = 1, 2, 3, \cdots$), the distribution of the expression

$$n^{m/2}[f\{S_n(x)\} - f\{\bar{V}_n(x)\}] \tag{5}$$

tends towards a function of bounded mean value and variance. For $m = 1$ it is a Gauss function with mean value 0 and finite variance. For any uneven m the distribution is symmetrical with respect to the zero point. If f is given, the limiting distribution *is essentially determined if in addition to $\bar{V}_n(x)$ one function of two variables, $\bar{U}_n(x, y)$, is known,*

$$\begin{aligned} \bar{U}_n(x, y) &= \frac{1}{n}\sum_{\nu=1}^{n} [V_\nu(x) - V_\nu(x)V_\nu(y)], & (x \leqq y) \\ &= \frac{1}{n}\sum_{\nu=1}^{n} [V_\nu(y) - V_\nu(x)V_\nu(y)], & (x \geqq y). \end{aligned} \tag{6}$$

For instance, in the case of the linear function ($m = 1$) defined in eq. (1), the (second order) variance of (5) is found as the Stieltjes integral

$$\int \psi(x)\psi(y)\, d\bar{U}_n(x, y) \tag{7}$$

and no mean values of higher order are required for computing the moments of any order, whatever m is.

For $m = 2$ the complete expression for the characteristic function of the asymptotic distribution of (5) is developed in Part III of this paper. It has the form

$$\frac{1}{D(ui)} \tag{8}$$

where $D(\lambda)$ is in general the Fredholm determinant of a symmetrical kernel that depends on the second derivative of $f\{V(x)\}$ at $V = \bar{V}_n$, on $\bar{V}_n$ and on $\bar{U}_n$. If the $V_\nu(x)$ are discontinuous distributions with saltus at k distinct points only, D is the determinant of a quadratic form of k variables. This happens to be the case with Pearson's χ^2 while the ω^2 distribution found by Smirnoff represents a fairly general case of the asymptotic distribution of second type.

PART I. PRELIMINARY THEOREMS

1. Asymptotically equal distributions. Let $K_1, K_2, K_3, \cdots$ be an infinite sequence of collectives, k_n the number of variables in K_n and A_n, B_n two functions of these variables, ($n = 1, 2, 3, \cdots$). The cumulative distribution functions of A_n and B_n will be denoted by $P_n(x)$ and $Q_n(x)$ respectively, i.e.

$$P_n(x) = \text{Prob}\,\{A_n \leqq x\}, \qquad Q_n(x) = \text{Prob}\,\{B_n \leqq x\} \tag{1}$$

and the expectation of $|A_n - B_n|$ by

$$E_n\{\,|A_n - B_n|\,\} \tag{2}$$

all these quantities being taken with respect to the distribution in K_n.

Two functions $F_n(x)$ and $G_n(x)$ both depending on the parameter n are said to be *asymptotically equal* if

$$\lim_{n=\infty} | F_n(x) - G_n(x) | = 0 \quad \text{uniformly in } x. \tag{3}$$

If this is the case for the cumulative distribution functions $P_n(x)$ and $Q_n(x)$ of A_n and B_n we shall also say that A_n and B_n have the same *asymptotic distribution*. Eq. (3) will also be written as $F_n(x) \sim G_n(x)$. The following can be proved:

LEMMA A. *If with increasing n the expectation of the absolute difference between A_n and B_n tends towards zero and if one of the functions $P_n(x)$ or $Q_n(x)$ is asymptotically equal to a function $F_n(x)$ that has a uniformly bounded derivative,* i.e.

$$\lim_{n=\infty} E_n\{| A_n - B_n |\} = 0, \qquad \left| \frac{dF_n(x)}{dx} \right| < M \quad \text{for all } n \tag{4}$$

then A_n and B_n have the same asymptotic distribution.

This statement, in a slightly different wording, was proved in an earlier paper [2] and the proof will not be repeated here. If one of the various definitions for "stochastical convergence" is used, one can also say that A_n and B_n, under the stated conditions, converge stochastically towards each other.

The Lemma A can be extended and modified in various ways. First, it is obvious that the expectation of $| A_n - B_n |$ can be replaced by that of any positive power $| A_n - B_n |^k$. With respect to F_n one could ask for the existence of a bounded derivative in all points except for a zero set only. Then P_n and Q_n would still converge everywhere except for this zero set and the definition of asymptotically equal distributions could be extended to this case. In the present paper this will not be done as it is not our purpose to strive for results of the possibly greatest generality.

2. Special class of statistical functions: quantics. Preliminary to the study of general statistical functions a special class which corresponds to quantics (homogeneous polynomials) of mth order must be discussed. Let $V_1(x)$, $V_2(x)$, $V_3(x)$, $\cdots$ be the cumulative distribution functions in a sequence of one-dimensional collectives C_1, C_2, C_3, $\cdots$ and $S_n(x)$ the repartition of a sample drawn from the n-dimensional collective K_n, with the distribution element

$$dV_1(x_1)dV_2(x_2) \cdots dV_n(x_n).$$

We introduce

$$T_n(x) = S_n(x) - \bar{V}_n(x), \qquad \bar{V}_n(x) = \frac{1}{n}\sum_{\nu=1}^{n} V_\nu(x). \tag{5}$$

Here, $nT_n(x)$ is obviously the *excess of observed values* $\leqq x$ over their expected number. Quantics of first, second, third, $\cdots$ order are then defined as

$$(6)\qquad \begin{aligned} f_1\{S_n(x)\} &= \int \psi(x)\, dT_n(x) \\ f_2\{S_n(x)\} &= \int\int \psi(x, y) dT_n(x) dT_n(y) \\ f_3\{S_n(x)\} &= \int\int\int \psi(x, y, z)\, dT_n(x)\, dT_n(y)\, dT_n(z) \end{aligned}$$

all integrals to be extended over the total range of x. Of course, only such ψ for which the respective integral exists are admitted. The first, f_1, is obviously a linear statistical function and the asymptotic distribution of $\sqrt{n}\, f_1$ is, under well-known conditions, a Gauss function with the mean value zero and the variance given in eq. (7) of the Introduction. In f_2, f_3, $\cdots$ the ψ may be supposed to be symmetrical with respect to their variables. It will be seen later (Part II, sec. 2) that the first derivative of f_2, the first and second derivatives of f_3, etc. vanish at the point $\bar{V}_n(x)$.

All the above functions f_1, f_2, f_3, $\cdots$ can be considered (if the ψ are continuous) as the limits of ordinary quantics in k variables. Choose k disjoint intervals I_1, I_2, $\cdots$, I_k on the x-axis, and call I_{k+1} their complement. Denote the increment of $V_\nu(x)$ within I_κ by $p_{\nu\kappa}$ and the increment of $S_n(x)$ by $\rho_{n\kappa}$. Obviously $p_{\nu\kappa}$ is the probability, within C_ν, of x falling in the interval I_k and $n\rho_{n\kappa}$ is the number of observed sample values in the same interval. We introduce the excess values ξ_κ:

$$(7)\qquad \xi_\kappa = \rho_{n\kappa} - \bar{p}_{n\kappa}, \qquad \bar{p}_{n\kappa} = \frac{1}{n}\sum_{\nu=1}^{n} p_{\nu\kappa}$$

and form the sums

$$(8)\qquad f_1 = \sum_{\kappa=1}^{k} \psi_\kappa \xi_\kappa, \qquad f_2 = \sum_{\kappa,\lambda}^{1\cdots k} \psi_{\kappa\lambda}\, \xi_\kappa \xi_\lambda, \qquad f_3 = \sum_{\kappa,\lambda,\mu}^{1\cdots k} \psi_{\kappa\lambda\mu}\, \xi_\kappa \xi_\lambda \xi_\mu, \cdots.$$

By selecting suitable sets of intervals I_1, I_2, $\cdots$, I_k and appropriate values for the constants ψ_κ, $\psi_{\kappa\lambda}$, $\cdots$, one can approximate the integrals (6) by sums of the form (8).

Our next task will be to find *asymptotic values for the expectation and for the moments of the quantities* defined in (8). Clearly a formula for the expectation of a power product $\xi_1^\alpha \xi_2^\beta \xi_3^\gamma \cdots$ where α, β, γ, $\cdots$ are positive integers, is the only thing we need. To arrive at such a formula we replace each of the one-dimensional collectives C_ν by a k-dimensional C_ν^* in the following way.

In C_ν^* the chance variable is a k-dimensional vector which can take $(k + 1)$ distinct values only: it can be zero or coincide with the unit vector parallel to one of the k axes. To the latter values of the variable we assign the probabilities $p_{\nu 1}, p_{\nu 2}, \cdots, p_{\nu k}$ and to the zero the probability

$$p_{\nu, k+1} = 1 - p_{\nu 1} - p_{\nu 2} - \cdots - p_{\nu k} \tag{9}$$

This quantity, of course, may vanish. The mean value of C_ν^* is the point with the coordinates $p_{\nu 1}, p_{\nu 2}, \cdots, p_{\nu k}$.

If the n collectives $C_1^*, C_2^*, \cdots, C_n^*$ are combined, the *sum* of the n observed vector values is a vector with the components $n\rho_{n1}, n\rho_{n2}, \cdots, n\rho_{nk}$. If in each C_ν^* the origin is shifted to the mean value and the coordinates with respect to the new origin are called $z_1, z_2, \cdots, z_k$, the sums of the observed $z_1, z_2, \cdots, z_k$-values will be $n\xi_1, n\xi_2, \cdots, n\xi_k$ rather than $n\rho_{n1}, n\rho_{n2}, \cdots, n\rho_{nk}$. Thus it is seen that all questions concerning the distributions of $\xi_1, \xi_2, \xi_3, \cdots$ can be answered on the basis of the well-known rules on the *addition of n independent chance variables*. This leads to the symbolic formula for the expectation:

$$E_n\{(n\xi_1)^\alpha (n\xi_2)^\beta (n\xi_3)^\gamma \cdots\} = \left(\sum_{\nu=1}^n Z_{\nu 1}\right)^\alpha \left(\sum_{\nu=1}^n Z_{\nu 2}\right)^\beta \left(\sum_{\nu=1}^n Z_{\nu 3}\right)^\gamma \cdots, \tag{10}$$

where on the right-hand side each term

$$Z_{\nu 1}^\kappa Z_{\nu 2}^\lambda Z_{\nu 3}^\mu \cdots \tag{11}$$

has to be replaced by

$$\int z_1^\kappa z_2^\lambda z_3^\mu \cdots dV_\nu^*(z). \tag{11'}$$

Here, obviously, $V_\nu^*(z)$ is the distribution function in C_ν^* and the expressions (11′) are in fact sums of $(k + 1)$ terms, for example

$$\begin{aligned} \int z_1 z_2 \, dV_\nu^*(z) &= p_{\nu 1}(1 - p_{\nu 1})(-p_{\nu 2}) + p_{\nu 2}(-p_{\nu 1})(1 - p_{\nu 2}) \\ &\quad + \sum_{\iota=3}^{k+1} p_{\nu\iota}(-p_{\nu 1})(-p_{\nu 2}) = -p_{\nu 1} p_{\nu 2}. \end{aligned} \tag{12}$$

It will be seen in the next section that only very few of these sums are needed for computing the asymptotic value of (10). Note that the value of (11′) can be expressed in terms of $p_{\nu 1}, p_{\nu 2}, p_{\nu 3}, \cdots$ alone if $\xi_1, \xi_2, \xi_3, \cdots$ only appear in the product.

3. Asymptotic expectation of excess-power products. We first consider the case where the sum of exponents $\alpha, \beta, \gamma, \cdots$ is an *even* number

$$\alpha + \beta + \gamma + \cdots = 2m. \tag{13}$$

On the right-hand side of (10) stands a sum of n^{2m} terms, each a product of $2m$ factors $Z_{\nu\kappa}$. It follows from (11′) that the absolute value of a product cannot surpass 1. The second subscripts are the same in each term: first α ones, then

β twos, γ threes, etc. The first subscripts are in each term a combination of $2m$ digits out of $\nu = 1, 2, 3, \cdots, n$. The number of those combinations which include s different ν-values, $(s = 1, 2, \cdots 2m)$, is

$$\binom{n}{s}K_s^{(m)} = \binom{n}{s}\left[s^{2m} - \binom{s}{1}(s-1)^{2m} + \cdots + \binom{s}{s-1}1^{2m}\right]. \tag{14}$$

Obviously, the $K_s^{(m)}$ are bounded (independent of n).

If $s > m$ the combination of first subscripts must include at least one ν-value that appears only once. All those products vanish since

$$\int z_\kappa dV_\nu^*(z_\kappa) = 0 \quad \text{for all } \kappa, \nu \tag{15}$$

due to the fact that the origin in the z-space coincides with the mean value of the distribution $V_\nu^*(z)$. Note that

$$\begin{aligned} \lim_{n\to\infty}\left[\binom{n}{s} : n^m\right] &= 0 && (s < m) \\ &= \frac{1}{m!} && (s = m). \end{aligned} \tag{16}$$

It follows that the sum of all terms in (10) that correspond to any $s < m$ are of the order $o(n^m)$ or smaller.

Thus, we arrive at an asymptotic expression for E_n by dividing both sides of (10) by n^m:

$$n^m E_n\{\xi_1^\alpha \xi_2^\beta \xi_3^\gamma \cdots\} \sim \frac{1}{n^m}\sum_\kappa \left(\prod_\nu Z_{\nu\kappa}\right) \tag{17}$$

where only such products on the right-hand side are retained which include *exactly m different ν-values each appearing twice.*

In analogy to (12) we compute

$$\begin{aligned} \int z_\iota z_\kappa dV_\nu^*(z) &= -p_{\nu\iota}p_{\nu\kappa} && (\iota \neq \kappa) \\ &= p_{\nu\iota}(1 - p_{\nu\iota}) && (\iota = \kappa) \end{aligned} \tag{18}$$

and write, for the sake of abbreviation

$$P_{\iota\kappa}^{(\nu)} = p_{\nu\iota}\delta_{\iota\kappa} - p_{\nu\iota}p_{\nu\kappa} = P_{\iota\kappa}^{(\nu)} \tag{19}$$

with the usual meaning of $\delta_{\iota\kappa}$ ($= 0$ if $\iota \neq \kappa$ and $= 1$ if $\iota = \kappa$). Then the sum to the right in (17) includes $(2m!)/2^m$ terms, each a product of m factors $P_{\iota\kappa}^{(\nu)}$. If each of the m couples ι, κ consists of two different figures, the respective product appears $\alpha!\,\beta!\,\gamma!\cdots$ times; if r couples are doubles $(\iota = \kappa)$ the multiplicity of the term is $2^{-r}\alpha!\,\beta!\,\gamma!\cdots$. Therefore, (17) takes the form

$$n^m E_n\{\xi_1^\alpha \xi_2^\beta \xi_3^\gamma \cdots\} \sim \frac{\alpha!\beta!\gamma!\cdots}{n^m}\sum_{\iota,\kappa,\nu} 2^{-r}P_{\iota_1\kappa_1}^{(\nu_1)}P_{\iota_2\kappa_2}^{(\nu_2)}\cdots P_{\iota_m\kappa_m}^{(\nu_m)}. \tag{20}$$

In this sum the upper indices are any set of m digits out of 1, 2, 3, $\cdots$, n and the subscripts are all sets of m couples including α ones, β twos, γ threes, etc. To each such set of m couples belong $\binom{n}{m}$ terms of the sum. The number of sets of couples is bounded (independent of n). The exponent r is the number of doubles ($\iota = \kappa$) among the m pairs.

The expression (20) admits of a transformation which renders it much more suitable. Assume that a set of couples ι, κ has been chosen according to the conditions and consider the product

$$\left(\sum_{\nu=1}^{n} P_{\iota_1\kappa_1}^{(\nu)}\right)\left(\sum_{\nu=1}^{n} P_{\iota_2\kappa_2}^{(\nu)} \cdots\right)\left(\sum_{\nu=1}^{n} P_{\iota_m\kappa_m}^{(\nu)}\right)2^{-r}. \tag{21}$$

Among the n^m terms which we obtain by developing (21) are all terms appearing in the sum (20), each of them repeated $m!$ times and, in addition,

$$n^m - \binom{n}{m}m! = n^m - n(n-1)(n-2)\cdots(n-m+1) \tag{22}$$

other products of m factors P. Since the difference (22) divided by n^m goes to zero with increasing n and each $|P|$ is smaller than 1, the additional terms have no importance. We therefore introduce the quantities

$$\bar{P}_{\iota\kappa} = \frac{1}{n}\sum_{\nu=1}^{n} P_{\iota\kappa}^{(\nu)} = \delta_{\iota\kappa}\text{-}\sum_{\nu=1}^{n} p_{\nu\iota} - \frac{1}{n}\sum_{\nu=1}^{n} p_{\nu\iota}\,p_{\nu\kappa}. \tag{23}$$

Then (20) can be written as

$$n^m E_n\{\xi_1^\alpha\,\xi_2^\beta\,\xi_3^\gamma\,\cdots\} \sim \frac{\alpha!\beta!\gamma!\cdots}{m!}\sum_{\iota,\kappa} 2^{-r}\bar{P}_{\iota_1\kappa_1}\bar{P}_{\iota_2\kappa_2}\cdots\bar{P}_{\iota_m\kappa_m}. \tag{24}$$

Here we have a sum of a finite number of terms. It will be supposed in all that follows that *the* $\bar{P}_{\iota\kappa}$ *as defined in* (23) *do not vanish identically as n increases indefinitely*.

Since in the sum (24) no upper indices appear, equal terms repeat themselves. We can, therefore, rearrange it, using the polynomial coefficients and absorbing at the same time the factor 2^{-r}. The final form of (24) is given in the following Lemma B_1, which also includes a statement for the case of an *uneven* sum of exponents $\alpha + \beta + \gamma + \cdots$. In fact, it is easily seen that if again half the sum is called m, no group of terms on the right-hand side of (10) exists that would supply a finite limit when divided by n^m. Thus we arrive at

LEMMA B_1. *If* $n\xi_\kappa$ *is the numerical excess of observed over expected quantities falling in the interval* I_κ, *the asymptotic expectation of the excess-power product* $\xi_1^\alpha\,\xi_2^\beta\,\xi_4^\gamma\cdots$ is given by

$$(\sqrt{n})^{\alpha+\beta+\gamma+\cdots}E_n\{\xi_1^\alpha\xi_2^\beta\xi_3^\gamma\cdots\} \sim 0 \qquad \textit{if } \alpha+\beta+\gamma+\cdots \textit{ uneven}$$

$$\sim \sum_\sigma \frac{\alpha!\beta!\gamma!\cdots}{\sigma_{11}!\sigma_{22}!\cdots\sigma_{12}!\cdots}(\tfrac{1}{2}\bar{P}_{11})^{\sigma_{11}}(\tfrac{1}{2}\bar{P}_{22})^{\sigma_{22}}\cdots\bar{P}_{12}^{\sigma_{12}}\bar{P}_{13}^{\sigma_{13}}\cdots, \tag{25}$$

$$\textit{if } \alpha+\beta+\gamma+\cdots \textit{ even}$$

the sum to be extended over all sets of non-negative integers $\sigma_{11}, \sigma_{22}, \cdots, \sigma_{12}, \cdots$ *that fulfill the conditions*

$$(25') \quad \sigma_{11} = \tfrac{1}{2}(\alpha - \sigma_{12} - \sigma_{13} - \cdots), \qquad \sigma_{22} = \tfrac{1}{2}(\beta - \sigma_{21} - \sigma_{23} - \cdots), \cdots$$

The $\overline{P}_{\iota\kappa}$ *as defined in* (23) *depend on two groups of mean values only, namely on*

$$(25'') \qquad \bar{p}_\kappa = \frac{1}{n}\sum_{\nu=1}^{n} p_{\nu\kappa} \quad \textit{and} \quad \overline{p_\iota p_\kappa} = \frac{1}{n}\sum_{\nu=1}^{n} p_{\nu\iota}\,p_{\nu\kappa}.$$

Some properties of the matrix $\overline{P}_{\iota\kappa}$ will be discussed in the next Section.

For practical computation, instead of (25), a *recursion formula* may be used which follows immediately from (24). Writing simply $(\alpha, \beta, \gamma, \cdots)$ for the sum in (24) the formula reads

$$(26) \qquad \begin{aligned}(\alpha, \beta, \gamma, \cdots) &= \tfrac{1}{2}(\alpha - 2, \beta, \gamma, \cdots)\overline{P}_{11} + \tfrac{1}{2}(\alpha, \beta - 2, \gamma, \cdots)\overline{P}_{22} + \cdots \\ &\quad + (\alpha - 1, \beta - 1, \gamma, \cdots)\overline{P}_{12} + (\alpha, \beta - 1, \gamma - 1), \cdots)\overline{P}_{23} + \cdots.\end{aligned}$$

If all the original distributions $V_\nu(x)$ are equal, this recursion formula, and from it (25), can be derived almost immediately from the theorem on the multiplication of characteristic functions with the addition of chance variables.

Note that the expectation of the product $\xi_\iota \xi_k$ is $\overline{P}_{\iota\kappa}/n$ for any value of n.

4. Asymptotic expectation and variance of quantics. We first state a characteristic property of the expression (25) for the expectation of an excess power product. Let us denote by $C_{\alpha,\beta,\gamma,\cdots}$ the right-hand side of (25) in the case of even $\alpha + \beta + \gamma + \cdots$. Then, if $C_{\alpha,\beta,\gamma\cdots}$ is expressed in terms of $\overline{P}_{\iota\kappa}$ and each time the subscript 2 is changed into 1, we arrive at the value of $C_{\alpha+\beta,o,\gamma,\cdots}$. This would not be the case if $C_{\alpha,\beta,\gamma,\cdots}$ were expressed in terms of p_ι, since e.g.

$$C_{11} = \overline{P}_{11} = \overline{p_1} - \overline{p_1 p_1}, \qquad C_{12} = \overline{P}_{12} = -\overline{p_1 p_2}.$$

In order to prove the statement we observe that the $C_{\alpha,\beta,\gamma,\cdots}$ can be derived from the coefficients in the development of the mth power of a **quadric:**

$$(27) \qquad \left(\tfrac{1}{2}\sum_{\iota,\kappa} \overline{P}_{\iota\kappa} t_\iota t_\kappa\right)^m = m!\sum \frac{C_{\alpha,\beta,\gamma,\cdots}}{\alpha!\beta!\gamma!\cdots} t_1^\alpha t_2^\beta t_3^\gamma \cdots.$$

It follows that

$$(27') \qquad C_{\alpha,\beta,\gamma,\cdots} = \frac{1}{m!}\frac{\partial^{2m}}{\partial t_1^\alpha\, \partial t_2^\beta\, \partial t_3^\gamma \cdots}\left[\left(\tfrac{1}{2}\sum_{\iota,\kappa} \overline{P}_{\iota\kappa} t_\iota t_\kappa\right)^m\right].$$

If in the subscripts of $\overline{P}_{\iota\kappa}$ the ones and twos are identified, the quadric becomes a function of $t_1 + t_2$, t_3, t_4, $\cdots$ and the derivative with respect to $\partial t_1^\alpha\, \partial t_2^\beta$ equals the derivative with respect to $\partial t_1^{\alpha+\beta}$. On the other hand, the latter derivative corresponds to the value of $C_{\alpha+\beta,0,\gamma,\cdots}$ in the form (27′).

Taking $m = 2$, $\alpha = \beta = \gamma = \delta = 1$, eq. (25) supplies

$$(28) \qquad n^2 E_n\{\xi_\iota \xi_\kappa \xi_\lambda \xi_\mu\} \sim \overline{P}_{\iota\kappa}\overline{P}_{\lambda\mu} + \overline{P}_{\iota\lambda}\overline{P}_{\kappa\mu} + \overline{P}_{\iota\mu}\overline{P}_{\kappa\lambda}.$$

According to the above statement this is correct whether ι, κ, λ, μ are or are not different from each other. Thus, if $\psi_{\iota\kappa\lambda\mu}$ is a symmetric set of constants, we have

$$n^2 E_n\{\sum_{\iota,\cdots\kappa} \psi_{\iota\kappa\lambda\mu}\xi_\iota\xi_\kappa\xi_\lambda\xi_\mu\} \sim 3\sum_{\iota\cdots\mu} \psi_{\iota\kappa\lambda\mu}\bar{P}_{\iota\kappa}\bar{P}_{\lambda\mu}. \tag{28'}$$

In general, the numerical factor to the right, i.e. the number of sets of couples drawn from $2m$ figures, is $(2m)!/2^m m! = 1\cdot 3 \cdots (2m-1)$. Thus we can state:

LEMMA B_2. *If a quantic f_{2m} is defined according to* (8) *with symmetric coefficients, its asymptotic expectation is given by*

$$n^m E_n\{f_{2m}\} \sim 1.3.5 \cdots (2m-1)\sum \psi_{\iota_1\iota_2\cdots\iota_{2m}}\bar{P}_{\iota_1\iota_2}\bar{P}_{\iota_3\iota_4}\cdots \bar{P}_{\iota_{2m-1}\iota_{2m}}. \tag{29}$$

Before applying this to the continuous case defined in (6), let us consider some characteristic properties of the matrix $\bar{P}_{\iota\kappa}$. According to the definition (19) of $P_{\iota\kappa}^{(\nu)}$ we have

$$\sum_{\iota,\kappa}^{1\cdots k} P_{\iota\kappa}^{(\nu)} t_\iota t_\kappa = \sum_{\iota=1}^{k} p_{\nu\iota} t_\iota^2 - \left(\sum_{\iota=1}^{k} p_{\nu\iota} t_\iota\right)^2 \tag{30}$$

and using (9) one easily derives from Schwarz' inequality

$$\Sigma p_{\nu\iota}t_\iota^2 - (\Sigma p_{\nu\iota}t_\iota)^2 \geqq p_{\nu,k+1}\Sigma p_{\nu\iota}t_\iota^2 .$$

Since $\bar{P}_{\iota\kappa}$ is the arithmetical mean of the $P_{\iota\kappa}^{(\nu)}$ it follows that the matrix $\bar{P}_\iota^{\ \kappa}$ is at least semi-definite and is positive definite except when all $p_{\nu,k+1} = 0$. In the latter case (if e.g. the k intervals cover the whole x-axis) one has

$$\sum_{\iota,\kappa}^{1\cdots k} \bar{P}_{\iota\kappa} = \frac{1}{n}\sum_{\nu=1}^{n}\left[\sum_{\iota=1}^{k} p_{\nu\iota} - \left(\sum_{\iota=1}^{k} p_{\nu\iota}\right)^2\right] = \frac{1}{n}\sum_{\nu=1}^{r} p_{\nu,k+1}(1-p_{\nu,k+1}) = 0 \tag{31}$$

which shows that here the reciprocal matrix $\bar{P}_{\iota\kappa}^*$ does not exist.

In the "complete" case, that is, with all $p_{\nu,k+1} = 0$, the elements in each horizontal or vertical line of the matrix $\bar{P}_{\iota\kappa}$ have the sum zero. It follows that the k homogenous equations $\Sigma\bar{P}_{\iota\kappa}x_\iota = 0$ have the solution $x_1 = x_2 = \cdots = x_k$ and, therefore, that the cofactors of all elements of $\bar{P}_{\iota\kappa}$ have one and the same value. For each single ν the determinant of $P_{\iota\kappa}^{(\nu)}$ can be computed:

$$| P_{\iota\kappa}^{(\nu)} | = p_{\nu1}p_{\nu2}\cdots p_{\nu k}p_{\nu,k+1}$$

If this is applied to the principal minors of the same determinant in the case $p_{\nu,k+1} = 0$, one finds the characteristic equation of the matrix $P_{\iota\kappa}^{(\nu)}$ to be

$$| \delta_{\iota\kappa} - \lambda P_{\iota\kappa}^{(\nu)} | = -\frac{d}{d\lambda}[(1-\lambda p_{\nu1})(1-\lambda p_{\nu2})\cdots(1-\lambda p_{\nu k})].$$

This shows that $(k-1)$ characteristic roots separate the abscissas $1/p_{\nu1}$, $1/p_{\nu2}$, $\cdots$, $1/p_{\nu k}$ (one root being zero).

The number k of intervals has nothing to do with the preceding argument leading to the eqs. (25) to (28). Also can the entire computation be repeated in terms of $dT_n(x_1)$, $dT_n(x_2)$, $dT_n(x_3)$, $\cdots$ instead of $\xi_1, \xi_2, \xi_3, \cdots$ if appropriate differentials are substituted for the $\bar{P}_{\iota\kappa}$. To find the latter ones we note that $p_{\nu\kappa}$ stands for the increment $dV_\nu(x)$. Thus, using $\delta(x, y)$ in analogy to $\delta_{\iota\kappa}$ (= 1 for $x = y$ and = 0 for $x \neq y$) we set

$$dU_\nu(x, y) = \delta(x, y)\, dV_\nu(x) - dV_\nu(x)\, dV_\nu(y) = \delta(x, y)\, dV_\nu(x) - dW_\nu(x, y) \tag{32}$$

which is equivalent to the definition of a function of 2 variables:

$$\begin{aligned} U_\nu(x, y) &= V_\nu(x) - V_\nu(x)V_\nu(y) = V_\nu(x) - W_\nu(x, y) && (x \leqq y) \\ &= V_\nu(y) - V_\nu(x)V_\nu(y) = V_\nu(y) - W_\nu(x, y) && (x \geqq y). \end{aligned} \tag{33}$$

Then $\bar{P}_{\iota\kappa}$ has to be replaced by

$$d\bar{U}_n(x, y) = \frac{1}{n} \sum_{\nu=1}^{n} dU_\nu(x, y) = \delta(x, y)\, d\bar{V}_n(x) - d\bar{W}_n(x, y). \tag{34}$$

This $d\bar{U}_n(x, y)$ is the expectation of $dT_n(x)\, dT_n(y)/n$.

The function

$$\bar{U}_n(x, y) = \frac{1}{n} \sum_{\nu=1}^{n} U_\nu(x, y) \tag{35}$$

is the difference of two cumulative distribution functions, one corresponding to a distribution along the straight line $x = y$ with the element $d\bar{V}_n(x)$ and another distribution over the whole plane with the element

$$d\bar{W}_n(x, y) = \frac{1}{n} \sum_{\nu=1}^{n} dV_\nu(x)\, dV_\nu(y). \tag{35'}$$

To each one-dimensional distribution $V_\nu(x)$ belongs one "*distribution excess*" $U_\nu(x, y)$ as defined in (33). The $\bar{P}_{\iota\kappa}^{(\nu)}$ are the increments of $U_\nu(x, y)$ within the product interval $dxdy$. It is seen from the preceding argument that the asymptotic moments of any quantic (6) or (8) depend only on the average $\bar{U}_n$ of the distribution excesses U_ν.

If a quantic is defined by (6) and the integrals on both sides exist, the asymptotic expectation of f_{2m} may be written in formal analogy to (29) as

$$n^m E_n\{f_{2m}\} \sim 1.3.5 \cdots (2m-1) \iint \cdots \int \psi(x_1, x_2, \cdots, x_{2m}) \times d\bar{U}_n(x_1, x_2)\, d\bar{U}_n(x_3, x_4) \cdots d\bar{U}_n(x_{2m-1}, x_{2m}). \tag{36}$$

This formula is identical with (29) if ψ has constant values in a finite number of intervals and vanishes outside these intervals. But it will be seen in the next section that (36) can be used in more general cases also.

For the sake of practical computation one may develop the righthand side of (36) into terms explicitly depending on the given averages $\bar{V}_n(x)$ and $\overline{W}_n(x, y)$. For example, in the case $m = 3$:

$$
\begin{aligned}
n^3E_n\{f_6\} \sim 1.3.5 \iiint & [\psi(x_1, x_1, x_2, x_2, x_3, x_3)\, d\bar{V}_n(x_1)\, d\bar{V}_n(x_2)\, d\bar{V}_n(x_3) \\
& - 3\psi(x_1, x_1, x_2, x_2, x_3, x_4)\, d\bar{V}_n(x_1)\, d\bar{V}_n(x_2)\, d\overline{W}_n(x_3, x_3) \\
& + 3\psi(x_1, x_1, x_2, x_3, x_4, x_5)\, d\bar{V}_n(x_1)\, d\overline{W}_n(x_2, x_3)\, d\overline{W}_n(x_4, x_5) \\
& - \psi(x_1, x_2, x_3, x_4, x_5, x_6)\, d\overline{W}_n(x_1, x_2)\, d\overline{W}_n(x_3, x_4)\, d\overline{W}_n(x_5, x_6)]
\end{aligned} \tag{37}
$$

In the general case, the numerical factors in the m-tuple integral are the binomial coefficients of order m.

The higher moments of quantics f_m can be computed in the same way as $E_n\{f_m\}$ since any power of f_m is a quantic again. The formulas, however, become more involved since the coefficients of f_m^s are not immediately given in a symmetric form. It will suffice to show here how the (second order) variance of f_2 can be found. The second moment is the expectation of

$$f_2^2 = \iiiint \psi(x, y)\psi(z, u)\, dT_n(x)\, dT_n(y)\, dT_n(z)\, dT_n(u). \tag{39}$$

Applying here eq. (28) we have

$$
\begin{aligned}
n^2E_n\{f_2^2\} \sim \iint \psi(x, y)\psi(z, u)[\, & d\bar{U}_n(x, y)\, d\bar{U}_n(z, u) \\
& + d\bar{U}_n(x, z)\, d\bar{U}_n(y, u) + d\bar{U}_n(x, u)\, d\bar{U}_n(y, z)].
\end{aligned} \tag{40}
$$

The first term in the brackets leads to the square of $n\, E_n\{f_2\}$ while the second and third terms, due to the symmetry of $\Psi(x, y)$, supply two equal integrals. Thus

$$
\begin{aligned}
&\operatorname{Var}\{nf_2\} \sim 2 \iint \psi(x, y)\psi(z, u)\, d\bar{U}_n(x, z)\, d\bar{U}_n(y, u) = \\
&2\Big[\iint \psi(x, x)\psi(y, y)\, d\bar{V}_n(x)\, d\bar{V}_n(y) - 2\iint \psi(x, y)\psi(y, z)\, d\bar{V}_n(y)\, d\overline{W}_n(x, z) \\
&\qquad + \iint \psi(x, y)\psi(z, u)\, d\overline{W}_n(x, z)\, d\overline{W}_n(y, u)\Big].
\end{aligned} \tag{41}
$$

In the same way moments and variances of any order can be computed for any quantic f_m.

5. Final statement on the limit of expectation of quantics. We shall prove the following:

LEMMA B_3. *Given a sequence of distributions $V_1(x)$, $V_2(x)$, $V_3(x)$, $\cdots$ and a quantic of order $2m$*

$$f_{2m} = \iint \cdots \int \psi(x_1, x_2, \cdots, x_{2m})\, dT_n(x_1)\, dT_n(x_2) \cdots dT_n(x_{2m})$$

assume that there exist a continuous function $\Psi(x)$ and a distribution $V(x)$ such that

$$\begin{aligned} &|\psi(x_1, x_2, \cdots x_{2m})| \leqq \Psi(x_1)\,\Psi(x_2) \cdots \Psi(x_{2m}) \\ &dV_\nu(x) \leqq dV(x) \text{ for } \quad |x| > X, \qquad \nu = 1, 2, 3, \cdots \end{aligned} \tag{42}$$

and that the integrals

$$\int \Psi^r(x)\, dV(x), \qquad (r = 1, 2, \cdots 2m), \tag{42'}$$

have finite values. Then, for any $\delta > 0$

$$\lim_{n=\infty} n^{m-\delta} E_n\{f_{2m}\} = 0. \tag{43}$$

This lemma, on which the main theorem of Part II is based, will be established if it is shown that the formula (36) holds true for functions ψ satisfying the conditions (42).

In the transition from the complete expression (10) for the expectation E_n to the asymptotic value (25) two essential steps were made. First, certain products of the form (11) have been omitted and, second, certain products of $P_{\iota\kappa}^{(\nu)}$ as defined in (19) have been arbitrarily added. This was allowed because each of the products was seen to be smaller than 1 and their number was of the order $O(n^{m-1})$. If a quantic in integral form (6) is considered which involves an infinite number of expressions like (10), a sharper estimate is necessary.

It is easily seen that each integral (11′) is a polynomial in $p_{\nu\kappa}$ including the product $p_{\nu1}p_{\nu2}p_{\nu3} \cdots$ and another factor which is certainly bounded whatever the $p_{\nu\kappa}$ are. Thus, if the expectation of $\xi_1\xi_2 \cdots \xi_{2m}$ is computed, each term of the form (11′) consists of a finite factor and the product $p_{\nu1}p_{\nu2} \cdots p_{\nu,2m}$. In passing to the expectation of the quantic, the $p_{\nu\kappa}$ have to be replaced by $dV_\nu(x_\kappa)$ and each neglected term in (10) leads to an expression like

$$\iint \cdots \int \psi(x_1, x_2, \cdots, x_{2m})\, dV_{\nu_1}(x_1)\, dV_{\nu_2}(x_2) \cdots dV_{\nu_\kappa}(x_\kappa). \tag{45}$$

According to the assumptions of B_3 this integral has a finite value. The number of neglected terms being of the order $O(n^{m-1})$ the omission of these terms is justified.

On the other hand, products of $P_{\iota\kappa}^{(\nu)}$ equal, except for the sign, products of $p_{\nu\iota}p_{\nu\kappa}$ as long as $\iota \neq \kappa$ and, except for a finite factor, products of $p_{\nu\iota}$ as often as $\iota = \kappa$. Again it is seen that the arbitrarily added terms sum up to integrals of the form (45). This shows that here too, if the conditions of B_3 are fulfilled, the procedure leading to (25) may be applied.

It follows that, under the conditions (42), if the integral (42′) has a finite value, eq. (36) is correct and (43) is an immediate consequence of it. On the other hand, it is obvious that weaker conditions than those given in B_3 would suffice to establish (43).

6. Theorem on products of n functions. The principal source of all explicit formulas on asymptotic distributions lies in certain properties of products of a great number of factors. Laplace devoted a part of his fundamental Treatise of Probability to these problems, but a complete outline of all results from a modern point of view is still lacking. In the third part of the present paper, a rather simple statement on this line will be used which may be formulated here as

LEMMA C. *Let $F_\nu(z_1, z_2, \cdots, z_k)$, $(\nu = 1, 2, 3, \cdots)$, be a sequence of analytic functions of k complex variables and G_n the product $F_1F_2 \cdots F_n$. Suppose that at the point $z_1 = z_2 = \cdots z_k = 0$ all F_ν have the value 1, vanishing first derivatives, and the second derivatives*

$$A_{\iota\kappa}^{(\nu)} = \frac{\partial^2 F_\nu}{\partial z_\iota \, \partial z_\kappa}. \tag{46}$$

Then

$$\lim_{n=\infty}\left[G_n\left(\frac{z_1}{\sqrt{n}}, \frac{z_2}{\sqrt{n}}, \cdots, \frac{z_n}{\sqrt{n}}\right) - \exp\left(\frac{1}{2n}\sum_{\iota,\kappa,\nu} A_{\iota\kappa}^{(\nu)} z_\iota z_\kappa\right)\right] = 0 \tag{47}$$

uniformly in each bounded region $|z_\iota| \leqq Z$ in which the absolute values of the third derivatives of all F_ν have an upper bound M.

In fact, the Taylor development of F_ν supplies under the conditions stated:

$$F_\nu(z_1, z_2, \cdots, z_k) = 1 + \tfrac{1}{2}\sum_{\iota,\kappa} A_{\iota\kappa}^{(\nu)} z_\iota z_\kappa + O(Z^3) \tag{48}$$

and, therefore,

$$\log F_\nu(z_1, z_2, \cdots, z_k) = \tfrac{1}{2}\sum_{\iota,\kappa} A_{\iota\kappa}^{(\nu)} z_\iota z_\kappa + O(Z^3). \tag{48′}$$

If here all z_ι are replaced by $z_\iota/\sqrt{n}$ and the equations added for $\nu = 1, 2, \cdots, n$ we obtain

$$\log G_n\left(\frac{z_1}{\sqrt{n}}, \frac{z_2}{\sqrt{n}}, \cdots, \frac{z_k}{\sqrt{n}}\right) = \frac{1}{2n}\sum_{\iota,\kappa,\nu} A_{\iota\kappa}^{(\nu)} z_\iota z_\kappa + nO\left(\frac{Z^3}{n\sqrt{n}}\right) \tag{49}$$

and this shows that the brackets on the left-hand side of (47) are $O(Z/\sqrt{n})$.— It is obvious that (47) would still hold if the condition concerning the third derivatives is replaced by a somewhat weaker one.

PART II. DIFFERENTIABLE STATISTICAL FUNCTIONS

1. Definitions. We consider a one-dimensional cumulative distribution function $V(x)$ as a point in the V-space. If two points $V_1(x)$ and $V_2(x)$ are given the functions

$$V_1(x) + t[V_2(x) - V_1(x)], \qquad 0 \leqq t \leqq 1 \tag{1}$$

represent the straight segment between $V_1(x)$ and $V_2(x)$. A subset of the V-space that includes all segments determined by its elements is called a *convex* domain.

Now, assume that a sequence of collectives with the distributions $V_1(x)$, $V_2(x)$, $V_3(x)$, $\cdots$ be given. We shall consider functions $f\{V(x)\}$ defined in a convex domain that includes particularly: (1) all average distributions $\bar{V}_n(x)$

$$\bar{V}_n(x) = \frac{1}{n} \sum_{\nu=1}^{n} V_\nu(x) \tag{2}$$

at least from a certain n on; (2) all repartitions $S_n(x)$ that can occur, i.e. the repartitions of n quantities that belong to the label sets of the given collectives (e.g. positive x, etc.). If $V^0(x)$ and $V(x)$ are any two points of the domain, the quantity

$$F(t) = f\{V^0(x) + t[V(x) - V^0(x)]\}, \qquad 0 \leqq t \leqq 1 \tag{3}$$

is a function of the real variable t. It will be supposed to admit derivatives with respect to t up to the order $r + 1$.

Following Volterra [9, 10] we define (in a slightly modified way) the derivative f' of a statistical function f in analogy to the set of partial derivatives of a function of several variables. If $V(x)$ would stand for a set of distinct variables $V_1, V_2, V_3, \cdots$ and $V^0(x)$ for their initial values $V_1^0, V_2^0, V_3^0, \cdots$ one would have

$$\frac{d}{dt} f\{V^0(x) + t[V(x) - V(x)]\}_{t=0} = \sum_y \frac{\partial f}{\partial V_y} (V_y - V_y^0)$$

where $\partial f/\partial V_y$ is the partial derivative of f with respect to V_y taken at the point $V_y = V_y^0$. Thus we write

$$\frac{d}{dt} f\{V^0(x) + t[V(x) - V^0(x)]\}_{t=0} = \int f'\{V^0(x), y\} d(V - V^0)(y) \tag{4}$$

and call f' which depends on $V^0(x)$ and on a scalar variable y, but not on $V(x)$, the (first) *derivative* of $f\{V(x)\}$ at the point $V^0(x)$. Only if a relation (4) is fulfilled for any two points of the convex domain, f is called a (one time) differentiable function.

The derivative of a linear function

$$A = \int \alpha(x)\, dV(x), \qquad B = \int \beta(x)\, dV(x), \cdots \tag{5}$$

is simply the factor $\alpha(y)$, $\beta(y)$ $\cdots$ respectively, independent of the point at which the derivative is taken. If f is given as a function of A, B, $\cdots$ one has

$$f'\{V(x), y\} = \frac{\partial f}{\partial A}\alpha(y) + \frac{\partial f}{\partial B}\beta(y) + \cdots. \tag{6}$$

The derivative of the non-linear function

$$f = \iint \psi(x, y)\, dV(x)\, dV(y) \tag{7}$$

is

$$f'\{V^0(x), y\} = \int [\psi(x, y) + \psi(y, x)]\, dV^0(x). \tag{8}$$

Note that an additive constant in f' (i.e. a quantity independent of y) has no significance since the integral of $d(V - V^0)$ vanishes. It follows from (6) that the first derivative of the mth order variance as defined in (2) of the Introduction, at the point $V^0(x)$ is

$$(y - a_0)^m - my\int (x - a_0)^{m-1}\, dV^0(x) \tag{9}$$

where a_0 is the mean value of $V^0(x)$.

In the same way derivatives of higher order can be introduced. The second derivative of $f\{V(x)\}$ is a function of $V^0(x)$, i.e. of the point at which the derivative is taken, and of two scalar variables y, z which correspond to the two subscripts in the case of a function of distinct variables. The definition of $f''\{V(x), y, z\}$ is given in the equation

$$\begin{aligned} &\frac{d^2}{dt^2} f\{V^0(x) + t[V(x) - V^0(x)]\}_{t=0} \\ &\qquad = \iint f''\{V^0(x), y, z\}\, d(V - V^0)(y)\, d(V - V^0)(z). \end{aligned} \tag{10}$$

The second derivative of a linear function is zero. The function (7) has the second derivative $\psi(z, y) + \psi(y, z)$ independently of $V^0(x)$. The mth order variance gives, twice differentiated

$$-2mz(y - a_0)^{m-1} + m(m - 1)yz \int (x - a_0)^{m-2}\, dV^0(x). \tag{11}$$

The variables y and z in f'' or in any additive term of f'' may be interchanged and a term depending on one of them may be added or omitted. Thus, f'' can always be written as a symmetric function of y, z without linear terms Accordingly, the second derivative of (7) is also $2\psi(y, z)$.

The derivative of rth order of f at the point $V^0(x)$ will be defined by the equation

$$\begin{aligned}(12)\quad &\frac{d^r}{dt^r} f\{V^0(x) + t[V(x) - V^0(x)]\}_{t=0} \\ &= \iint \cdots \int f^{(r)}\{V^0(x), y_1, y_2, \cdots, y_r\}\, d(V - V^0)(y_1) \cdots d(V - V^0)(y_r).\end{aligned}$$

Here, for given $V^0(x)$, $f^{(r)}$ may be supposed to be a symmetric function of the r variables $y_1, y_2, \cdots, y_r$. The rth derivative of the mth order variance is

$$\begin{aligned}(13)\quad &\frac{(-1)^r m!}{(m - r + 1)!} y_1 y_2 \cdots y_r \\ &\times \left[(m - r + 1) \int (x - a_0)^{m-r}\, dV^0(x) - \sum_{\kappa=1}^{r} \frac{(y_\kappa - a_0)^{m-r+1}}{y_\kappa} \right].\end{aligned}$$

In the case $r = m$ the expression becomes independent of $V^0(x)$, viz.

$$(13') \qquad (-1)^m m!\, y_1 y_2 \cdots y_m (1 - m)$$

where terms depending on less than r of the variables $y_1, y_2, \cdots, y_r$ have been omitted.

If the definitions (4), (10), (12) are confronted one can see that $f''\{V, y, z\}$ is the first derivative of $f'\{V, y\}$ etc. For proofs see [9] and [10].

2. Taylor development. The function $F(t)$ defined in (3) admits the development

$$(14) \quad F(1) - F(0) = F'(0) + \frac{1}{2!} F''(0) + \cdots + \frac{1}{r!} F^{(r)}(0) + \frac{1}{(r + 1)!} F^{(r+1)}(\vartheta)$$

where ϑ is some quantity between zero and one. According to (3) the left-hand side equals the difference $f\{V(x)\} - f\{V^0(x)\}$. The expressions $F'(0), F''(0), \cdots, F^{(r)}(0)$ are the derivatives as defined in eqs. (4), (10), (12). In the last term to the right, one has to introduce the distribution

$$(15) \qquad V'(x) = V^0(x) + \vartheta[V(x) - V^0(x)]$$

and then to take the $(r + 1)$st derivative of f at the point $V'(x)$.

For a given $V^0(x)$ each one of the terms on the right-hand side of (14) is a function of $V(x)$. Except for the last one—in which ϑ depends in a certain way on $V(x)$—they are *quantics* with respect to $V(x) - V^0(x)$, of the same kind as those considered in Part I. (There we had S_n instead of V and $\bar{V}_n$ instead of V^0).

The rth term of (14) can be written as

$$F_r = \frac{1}{r!} \iint \cdots \int \psi(x_1, x_2, \cdots, x_r)\, d(V - V^0)(x_1) \cdots d(V - V^0)(x_r) \tag{16}$$

where according to (12)

$$\psi(x_1, x_2, \cdots, x_r) = f^{(r)}\{V^0(x), x_1, x_2, \cdots, x_r\}. \tag{16'}$$

To find the characteristic properties of F_r we compute its derivatives at a point $V_1(x)$. To do this we must replace in (16) the $V(x)$ by

$$V_1(x) + t[V(x) - V_1(x)]$$

then differentiate the product

$$\prod_{\kappa=1}^{r} d[(V_1 - V^0)(x_\kappa) + t(V - V_1)(x_\kappa)] \tag{17}$$

with respect to t, and finally set $t = 0$. The derivative consists of r terms the first of which will be

$$d(V - V_1)(x_1) \prod_{\kappa=2}^{r} d(V_1 - V^0)(x_\kappa).$$

Due to the fact that ψ may be supposed as a symmetric function, all r terms supply the same integral. Thus the derivative of F_r with respect to t at the point $t = 0$ can be written as

$$\frac{1}{(r-1)!} \iint \cdots \int \psi(x_1, x_2, \cdots, x_r)\, d(V - V_1)(x_1) \prod_{\kappa=2}^{r} d(V_1 - V^0)(x_\kappa).$$

Comparing this with the formula (4) which defines the first derivative of a statistical function and writing y instead of x and $V(x)$ instead of $V_1(x)$, we find

$$\begin{aligned} &F_r'\{V(x), y\} = \\ &\frac{1}{(r-1)!} \iint \cdots \int \psi(y, x_2, x_3, \cdots, x_r)\, d(V - V^0)(x_2) \cdots d(V - V^0)(x_r). \end{aligned} \tag{18}$$

This is the first derivative of $F_r\{V(x)\}$ at the point $V(x)$. It vanishes at the point $V(x) = V^0(x)$.

The integral in (18) has the same form as that in (14) except that its multiplicity is $(r - 1)$ rather than r. Thus it is immediately seen how the higher derivatives of F_r can be found. For the second derivative $F_r''\{V(x), y, z\}$ we have simply to replace $(r - 1)!$ in (18) by $(r - 2)!$, then x_2 by z and finally to omit in the product the differential $d(V - V^0)(x_2)$. This procedure can be continued up to the derivative of order $(r - 1)$. The rth derivative, finally, will be

$$F_r^{(r)}\{V(x), y_1, y_2, \cdots, y_r\} = \psi(y_1, y_2, \cdots, y_r) \tag{19}$$

independent of $V(x)$ and, according to (16′), equal to the rth derivative of $f\{V(x)\}$ at the point $V^0(x)$. It is also seen that all integrals of the form(16) or (18) vanish if $V(x)$ equals $V^0(x)$. The results can be summarized as follows: The sth term, $(s = 1, 2, \cdots r)$, of the development (14) is a function of $V(x)$ for which all derivatives at the point $V^0(x)$ except that of order s vanish while this one equals the sth derivative of the original function $f\{V(x)\}$ at $V^0(x)$. The complete analogy of (14) with the Taylor development of a function of distinct variables is thus evident.

If we assume that $f\{V(x)\}$ is a function whose first $(r - 1)$ derivatives vanish at the point $V^0(x)$, eq. (14) takes the form

$$\begin{aligned} V(x) - V^0(x) = \frac{1}{r!}\int\int\cdots\int f^{(r)}\{V^0(x), y_1, y_2, \cdots, y_r\} \\ \cdot d(V - V^0)(y_1)\cdots d(V - V^0)(y_r) \\ + \frac{1}{(r+1)!}\int\int\cdots\int f^{(r+1)}\{V'(x), y_1, y_2, \cdots, y_{r+1}\} \\ \cdot d(V - V^0)(y_1)\cdots d(V - V^0)(y_{r+1}). \end{aligned} \tag{20}$$

By applying to this formula the lemmas A and B of Part I, we shall arrive at the general theorem on asymptotic distributions that is the principal goal of this paper.

3. General theorem. The main result to be derived in the general theory of asymptotic distributions is that the so-called normal distribution represents the first element in an infinite sequence which includes the asymptotic distributions of all differentiable statistical functions, except certain irregular cases. The Gauss distribution covers in fact only those functions whose Taylor development starts with the first (linear) term, in particular the linear statistical functions themselves. If the first $(r - 1)$ terms in the development vanish, the asymptotic distribution of type r becomes valid.

THEOREM I: *Let $V_1(x), V_2(x), V_3(x), \cdots$ be an infinite sequence of distributions and $f\{V(x)\}$ a statistical function with derivatives up to order $(r+1)$. Denote by $S_n(x)$ the repartition of the n label values in the collective with the distribution element $dV_1(x)$, $dV_2(x) \cdots dV_n(x)$ and by $\bar{V}_n(x)$ the arithmetical mean of $V_1(x)$, $V_2(x), \cdots, V_n(x)$. If for large n the first $(r - 1)$ derivatives of $f\{V(x)\}$ at the point $\bar{V}_n(x)$ vanish and the rth derivative equals $\psi_n(y_1, y_2, \cdots, y_r)$, then the distribution of*

$$A_n = n^{r/2}[f\{S_n(x)\} - f\{\bar{V}_n(x)\}] \tag{21}$$

is asymptotically equal to the distribution of the rth order quantic

$$B_n = \frac{n^{r/2}}{r!} \iint \cdots \int \psi_n(x_1, x_2, \cdots, x_r) \cdot d(S_n - \bar{V}_n)(x_1)\, d(S_n - \bar{V}_n)(x_2) \cdots d(S_n - \bar{V}_n)(x_r) \tag{22}$$

under the following conditions:

a) *The distribution of* (22) *has a uniformly bounded derivative for all* n;

b) *Within a convex domain in the V-space that includes all* $\bar{V}_n(x)$ *from a certain* n *on, and all* $S_n(x)$ *that can occur, the* $(r+1)$*st derivative of* $f\{V(x)\}$ *is smaller in absolute value than a product* $\Psi(y_1)\Psi(y_2) \cdots \Psi(y_{r+1})$ *whereby the integrals* $\int [\Psi(x)]^k\, dV_\nu(x)$ *for* $k = 1, 2, \cdots, 2(r+1)$ *have a finite upper bound for* $\nu = 1, 2, 3, \cdots$.

In order to prove this we introduce in eq. (20) $S_n(x)$ for $V(x)$ and $\bar{V}_n(x)$ for $V^0(x)$, and multiply both sides by $n^{r/2}$. Using the notations (21) and (2) and writing T_n for $(S_n - \bar{V}_n)$, the equation reads

$$A_n - B_n = \frac{n^{r/2}}{(r+1)!} \iint \cdots \int f^{(r+1)}\{V'(x), y_1, y_2, \cdots, y_{r+1}\}\, dT_n(y_1) \cdots dT_n(y_{r+1}). \tag{32}$$

According to Lemma A the theorem will be verified if we can show that the expectation of the absolute value of the right-hand expression in (23) tends to zero.

According to the Schwarz inequality one has, for any real C:

$$E_n\{|C|\} \leqq \sqrt{E_n\{C^2\}}. \tag{24}$$

For fixed values of $\bar{V}_n$ and S_n the integral on the right-hand side of (23) is a quantic of order $(r+1)$ with the coefficients $\psi_{r+1}(y_1, y_2, \cdots, y_{r+1})$. The square of this integral is a quantic of order $2(r+1)$ whose coefficients are a finite number (depending only on r) of terms each of which is a product of two ψ_{r+1}-values implying $2(r+1)$ variables $y_1, y_2, \cdots, y_{2(r+1)}$. The absolute value of these coefficients is, therefore, according to the condition b) smaller than a finite factor times the product $\Psi(y_1)\ \Psi(y_2) \cdots \Psi(y_{2(r+1)})$ and thus fulfills the condition of lemma B_3. If the right-hand side of (23) is identified with C, the expectation of C^2 is, except for a finite factor, the product of n^r times the expectation of the above-mentioned quantic of order $2(r+1)$. It then follows from lemma B_3 that the limit of $E_n\{C^2\}$ is zero and from (24):

$$\lim_{n=\infty} E_n\{|C_n|\} = \lim_{n=\infty} E_n\{|A_n - B_n|\} = 0.$$

This accomplishes the proof of Theorem I.

If we apply here what was shown in Part I about the asymptotic distribution of a quantic, we can also state the following.

THEOREM II: *Under the conditions of Theorem I, the asymptotic distribution of a differentiable statistical function* $f\{S_n(x)\}$ *is essentially determined by*

a) *the average distribution* $\bar{V}_n(x)$;

b) *the first non-vanishing derivative of* $f\{V(x)\}$ *at the point* $\bar{V}_n(x)$;

c) *the average distribution excess*

$$\begin{aligned}\bar{U}_n(x, y) &= \bar{V}_n(x) - \frac{1}{n}\sum_{\nu=1}^{n} V_\nu(x)V_\nu(y), & x \leqq y \\ &= \bar{V}_n(y) - \frac{1}{n}\sum_{\nu=1}^{n} V_\nu(x)V_\nu(y), & x \geqq y.\end{aligned} \tag{25}$$

By "essentially determined" is meant determined except for an additional function whose moments of any order are zero. The statement then follows from Theorem I in connection with the fact that the asymptotic moments of quantics have been computed in Part I from the values of $\bar{U}_n(x, y)$.

That functions with all moments vanishing exist has been known for a long time. A simple example given by Shohat and Tamarkin [6] is the following. Let κ be a positive constant smaller than $\frac{1}{2}$, and $u = x^\kappa$, $k = \tan \kappa\pi$. Then, the density (positive or negative)

$$\varphi(x) = e^{-u} \sin (ku) = Im\, e^{-u(1-ki)} \tag{26}$$

fulfills the condition. In fact, the nth moment of (26) is the (vanishing) imaginary part of the integral

$$\frac{1}{\kappa}\int_0^\infty u^{(n+1/\kappa)-1} e^{-u(1-ki)}\, du = \frac{(-1)^{n-1}}{\kappa} (\cos \kappa\pi)^{(n+1/\kappa)} \Gamma\left(\frac{n+1}{\kappa}\right). \tag{27}$$

Since $\varphi(x)$ takes negative values of the amount e^{-u} it can be superimposed to a given distribution density only in cases where the original density remains greater than some multiple of $e^{-u} = \exp(-x^k)$. It can be shown that the moment problem is determinate (i.e. the distribution determined by the moments in a unique way) if the density vanishes at infinity at a sufficiently strong degree.

From the standpoint of statistical theory two distributions with the same moments throughout may be considered as equivalent. This justifies the terminology used in Theorem II. On the other hand, Theorem I is independent of this restriction: The asymptotic distribution of the statistical function $f\{S_n(x)\}$ is under the given conditions identical with that of the corresponding quantic of mth order. A detailed discussion of the case $m = 2$ will be given in Part III. Here follow some illustrations for the general case.

4. Illustrations. The existence of asymptotic distributions of higher types can be exemplified in a comparatively simple way if we start from any known asymptotic distribution of a statistical function.

Let us assume that $g\{V(x)\}$ is a function fulfilling the condition

$$g\{\bar{V}_n(x)\} = 0 \tag{28}$$

for all n, and that the asymptotic c.d.f. for $g\{S_n(x)\}$ is known. There will be some positive integer r such that

(29) $$\text{Prob}\,[g\{S_n(x)\} \leqq zn^{-r/2}] \sim \Phi_n(z).$$

If, for instance, g is a linear statistical function r will be 1 and, under well-known conditions, $\Phi_n(x)$ a normal (Gaussian) c.d.f. with finite variance depending on n.

Now, let f be an ordinary function of g and thus another statistical function which may be denoted by $f\{V(x)\}$. According to the rules of differentiation we have

(30) $$f'\{V(x), y\} = \frac{df}{dg}\, g'\{V(x), y\}$$

and analogous relations can be derived for the derivatives of higher order. In particular, the following statement, valid in ordinary differential calculus, holds true: If $g\{V(x)\}$ has derivatives of every order and if the first s derivatives of f with respect to g vanish at some point $g = g\{V_1(x)\}$ then also the s first derivatives of f with respect to $V(x)$ will be zero at $V(x) = V_1(x)$. In this way we can devise statistical functions, with vanishing derivatives, for which the asymptotic distribution is known.

For the sake of simplicity we may assume that (29) holds with $r = 1$ and that $f(g)$ is a monotonic increasing function, given in the form

(31) $$f(g) = g^s[1 + \alpha(g)]$$

with s a positive integer, and the inverse function

(31′) $$g(f) = f^{1/s}[1 + \beta(f)]$$

where $\beta(f)$ goes to zero with $f \to 0$. Then, from (29):

(32) $$\text{Prob}\,[f\{S_n(x)\} \leqq zn^{-(s/2)}] \sim \Phi_n(z')$$

if z and z' are connected by

$$n^{-\frac{1}{2}}z' = g(n^{-(s/2)}z) = n^{-\frac{1}{2}}z^{1/s}[1 + \beta(n^{-(s/2)}z)].$$

It follows that

$$z' - z^{1/s} \sim 0$$

and if $\Phi_n(z')$ is supposed to be continuous, (32) becomes

(33) $$\text{Prob}\,[f\{S_n(x)\} \leqq zn^{-(s/2)}] \sim \Phi_n(z^{1/s}).$$

This is a distribution of type s.

Take as an example for g the arithmetical mean

(34) $$g\{S_n(x)\} = \frac{x_1 + x_2 + \cdots + x_n}{n} - \bar{a}_n$$

where $x_1, x_2, \cdots, x_n$ are the observed values and $\bar{a}_n$ is the arithmetical mean of the mean values of $V_\nu(x)$. Then, under certain restrictions for the $V_\nu(x)$, there exists a bounded sequence h_n^2 so that

$$\text{Prob}[\sqrt{n}\, g \leqq z] \sim \Phi_n(z) = \frac{h_n}{\sqrt{\pi}} \int_{-\infty}^{z} e^{-h_n^2 u^2}\, du.$$

Now if we choose

$$f = 6(g - \sin g) = g^3\left(1 - \frac{g^2}{20} + \cdots\right)$$

the asymptotic distribution of f will be given by

$$\text{Prob}\,[n\sqrt{n} f \leqq z] \sim \Phi_n(\sqrt[3]{z}) = \frac{h_n}{\sqrt{\pi}} \int_{-\infty}^{z^{1/3}} e^{-h_n^2 u^2}\, du$$

with the probability density

$$\frac{h_n}{3\sqrt{\pi}}\, z^{-(2/3)}\, e^{-h_n^2 z^{2/3}}.$$

Similar examples can be drawn from the asymptotic distribution of $n\chi^2$ if one asks for the distribution of appropriate functions of $n\chi^2$, etc.

PART III. SECOND-TYPE ASYMPTOTIC DISTRIBUTION

1. Statement of the problem. We now propose to study the asymptotic distribution of a quantic of second order as defined in eq. (6) of Part I. It has been shown in Part II that this covers the case of any statistical function of which the first but not the second derivative at the critical point vanishes.

Independently of what was said before, the problem can be stated in the following way. Given a function $\psi(x, y)$ and a sequence of cumulative distribution functions $V_1(x), V_2(x), V_3(x) \cdots$. Let $\bar{V}_n(x)$ be the arithmetical mean of $V_1(x), V_2(x), \cdots, V_n(x)$ and $S_n(x)$ the repartition of a sample $z_1, z_2, \cdots, z_n$ drawn from the collective with the distribution element $dV_1(z_1)\, dV_2(z_2), \cdots, dV_n(z_n)$, that is: $nS_n(x)$ is the number of those of the observed values $z_1, z_2, \cdots, z_n$ that are smaller than or equal to x. Then the quantity

$$f = \iint \psi(x, y)\, dT_n(x)\, dT_n(y), \quad \text{where } T_n(x) = S_n(x) - \bar{V}_n(x) \tag{1}$$

is determined by the observations $z_1, z_2, \cdots, z_n$. We ask for the distribution of f at large values of n.

Without loss of generality, the function $\psi(x, y)$ can be supposed to be symmetrical. If, in particular, $\psi(x, y) = \psi(x)\psi(y)$, the quantity f becomes the square of

$$\int \psi(x)\, dT_n(x) = \frac{1}{n} \sum_{\nu=1}^{n} \left[\psi(z_\nu) - \int \psi(z)\, dV_\nu(x)\right] \tag{2}$$

and its asymptotic distribution can be computed in the manner shown in the last section of Part I. Another example would be

$$\psi(x, y) = g(x) \qquad (x \leqq y) \tag{3}$$
$$= g(y) \qquad (x \geqq y).$$

In this case, integration by parts shows that

$$f\{S_n(x)\} = \int g'(x) T_n^2(x)\, dx \tag{4}$$

where g' is the derivative of g. This is the statistical function that takes the place of χ^2 in continuous problems. See Introduction eq. (3).

Note that the "excess" $T_n(x)$ vanishes at $x = \pm\infty$ and that for sufficiently large x the increment $dT_n(x)$ equals $-d\bar{V}_n(x)$. Thus, conditions for the existence of the integrals in (1), (2), (4), etc. can be expressed in terms of the given functions $\psi(x, y)$ and $V_\nu(x)$.

We shall first study the special case that implies so-called discontinuous chance variables. In our terminology it is the function $\psi(x, y)$ that has to be specified. Let $I_1, I_2, \cdots, I_k$ be k mutually exclusive one-dimensional intervals (or groups of intervals) and I_{k+1} their complement. Assume that $\psi(x, y)$ has a constant value when x falls in I_ι and y falls in I_κ, $(\iota, \kappa = 1, 2, \cdots, k+1)$. The increments of $S_n(x)$, $\bar{V}_n(x)$, $T_n(x)$ in the interval I_κ will be called ρ_κ, $\bar{p}_\kappa$, ξ_κ respectively. Clearly, $n\rho_\kappa$ is the number of observed values falling in I_κ, $n\bar{p}_\kappa$ is the expected number of such values, and $n(\rho_\kappa - \bar{p}_\kappa) = n\xi_\kappa$ the excess of observed over expected numbers. Note that the given distributions $V_\nu(x)$ determine increments $p_{\nu\kappa}$ in the interval I_κ and that

$$\bar{p}_\kappa = \frac{1}{n}(p_{1\kappa} + p_{2\kappa} + \cdots + p_{n\kappa}). \tag{5}$$

Since the sum of all ξ_κ must be zero we can replace ξ_{k+1} by

$$\xi_{k+1} = -\xi_1 - \xi_2 - \cdots - \xi_k. \tag{6}$$

Thus, the integral (1) can now be written as a sum of k^2 terms

$$f\{S_n(x)\} = \sum_{\iota,\kappa}^{1\cdots k} \psi_{\iota\kappa}\xi_\iota\xi_\kappa \tag{7}$$

like that introduced in the second eq. (8) of Part I.

Our next task will be to find the asymptotic distribution of (7) which depends on the matrix $\psi_{\iota\kappa}$, $(\iota, \kappa = 1, 2, \cdots, k)$, and on the succession of probability values $p_{\nu\kappa}$, $(\nu = 1, 2, 3, \cdots; \kappa = 1, 2, \cdots k)$. The matrix $\psi_{\iota\kappa}$ in k variables will be supposed to be symmetrical.

2. Characteristic function. We define our chance variable as

$$x = \frac{n}{2} f. \tag{8}$$

All summations, here and in what follows, are to be extended from 1 to k if not otherwise indicated. If $P_n(x)$ is the c.d.f. of x, that is

$$\text{Prob}\left\{\frac{n}{2} f \leqq x\right\} = P_n(x) \tag{9}$$

the characteristic function (c.f.) is defined by

$$Q_n(u) = E\{e^{xui}\} = \int e^{xui} dP_n(x). \tag{10}$$

In order to compute Q_n we assume that the quadratic form (8) is transformed, by a linear transformation, into a sum of squares. Using appropriate (in general complex) coefficients $\alpha_{\iota\kappa}$ one can write

$$x = \frac{n}{2} (\eta_1^2 + \eta_2^2 + \cdots + \eta_k^2), \qquad \eta_\iota = \sum_\kappa \alpha_{\iota\kappa} \xi_\kappa . \tag{11}$$

(The form $\psi_{\iota\kappa}$ is here supposed to be non-singular which, however, means no loss of generality). It will be seen later that explicit knowledge of the $\alpha_{\iota\kappa}$ is not needed.

Now, for any real or complex y, the identity holds:

$$e^{\frac{1}{2}y^2} = \frac{1}{\sqrt{2\pi}} \int e^{\frac{1}{2}t^2+yt}\, dt. \tag{12}$$

If we write v for $\sqrt{ui}$ and replace in (12) successively y by $v\sqrt{n}\, \eta_1$, $v\sqrt{n}\, \eta_2$, $\cdots$ we find

$$e^{xui} = (2\pi)^{-k/2} \iint \cdots \int \exp\left[-\tfrac{1}{2}\sum t_\kappa^2 + v\sqrt{n}\sum z_\kappa \xi_\kappa\right] dt_1\, dt_2 \cdots dt_k \tag{13}$$

where

$$\sum z_\kappa \xi_\kappa = \sum \eta_\kappa t_\kappa, \qquad z_\kappa = \sum_\iota \alpha_{\iota\kappa} t_\iota, \qquad (\kappa = 1, 2, \cdots, k). \tag{14}$$

Since the first exponential factor in the integrand is a constant with respect to the chance variable, the expected value of e^{xui} is given by

$$Q_n(u) = E\{e^{xui}\} = (2\pi)^{-k/2} \iint \cdots \int \exp\left[-\tfrac{1}{2}\sum t_\kappa^2\right] G_n\, dt_1\, dt_2 \cdots dt_k \tag{15}$$

with

$$G_n = E\{\exp[v\sqrt{n} \sum z_\kappa \xi_\kappa]\}. \tag{16}$$

In order to find G_n we consider the following n collectives C_1, C_2, $\cdots$, C_n with discontinuous, $(k + 1)$-valued distributions: In C_ν the label values are z_1, z_2, $\cdots$, z_k, and z_{k+1}, with $z_{k+1} = 0$, their probabilities $p_{\nu 1}$, $p_{\nu 2}$, $\cdots$, $p_{\nu,k+1}$. The c.f. of this distribution at the point $-iv/\sqrt{n}$ is

$$\sum_{\kappa=1}^{k+1} p_{\nu\kappa} e^{vz_\kappa/\sqrt{n}}. \tag{17}$$

If we multiply the n expressions (17) for $\nu = 1, 2, \cdots n$ the product will be—according to well-known rules of probability calculus—the c.f. for the distribution of the *sum* of the n label components in the collective formed by combining $C_1, C_2, \cdots, C_n$. This sum is

$$\sum n p_\kappa z_\kappa$$

and therefore,

$$E\left\{\exp\left[\frac{v}{\sqrt{n}} \sum n p_\kappa z_\kappa\right]\right\} = \prod_{\nu=1}^{n}\left[\sum_{\kappa=1}^{k+1} p_{\nu\kappa} e^{v z_\kappa/\sqrt{n}}\right]. \tag{18}$$

Multiplying both sides of this equation by

$$\exp\left[-\frac{v}{\sqrt{n}} \sum n \bar{p}_\kappa z_\kappa\right] = \exp\left[-\sum_{\nu=1}^{n} \frac{v}{\sqrt{n}} \sum_\kappa p_{\nu\kappa} z_\kappa\right] \tag{19}$$

and using the abbreviation

$$\bar{z}_\nu = \sum_\kappa p_{\nu\kappa} z_\kappa \tag{20}$$

we arrive at

$$G_n = E\{\exp[v\sqrt{n} \sum \xi_\kappa z_\kappa]\} = F_1 F_2 \cdots F_n \tag{21}$$

with

$$F_\nu = \sum_{\kappa=1}^{k+1} p_{\nu\kappa} e^{v(z_\kappa - \bar{z}_\nu)/\sqrt{n}}. \tag{22}$$

This solves the problem: By inserting (21), (22) in (15) and carrying out the integration with respect to $t_1, t_2, \cdots, t_k$ one has expressed $Q_n(u)$ in terms of the given $p_{\nu\kappa}$ and of the coefficients $\alpha_{\iota\kappa}$ which link the z_κ to the t_κ. This expression for $Q_n(u)$ holds for all n.

We have still to show that the integral (15) exists, at least for small $|u|$ or $|v|$, independently of the value of n. For this purpose we develop F_ν, as given in (22), in the neighborhood of $v = 0$. At this point $F_\nu = 1$ and the first derivative vanishes by virtue of (20). We thus have

$$F_\nu = 1 + \frac{v^2}{2n} \sum_{\kappa=1}^{k+1} p_{\nu\kappa}(z_\kappa - \bar{z}_\nu)^2 e^{\vartheta_\kappa v(z_\kappa - \bar{z}_\nu)/\sqrt{n}} \tag{23}$$

with $|\vartheta_\kappa| \leqq 1$. From the definition of z_κ in (14) it follows that the ratio $|z_\kappa|/T$ with

$$T^2 = t_1^2 + t_2^2 + \cdots + t_k^2$$

has an upper bound depending on the $\alpha_{\iota\kappa}$ only. On the other hand, according to (20), $\bar{z}_\nu$ is a weighted mean of the z_κ and, therefore, $|z_\kappa - \bar{z}_\nu|$ will not surpass twice the maximum $|z_\kappa|$:

$$|z_\kappa - \bar{z}_\nu| \dot{<} \alpha T \tag{25}$$

where α is a positive function of the coefficients $\alpha_{\iota\kappa}$ which, in turn, are determined by the $\psi_{\iota\kappa}$. Introducing (25) in (23) we find

$$| F_\nu | < 1 + \frac{| v |^2 \alpha^2 T^2}{2n} e^{|v|\alpha T/\sqrt{n}} \leqq e^{|v^2|\alpha^2 T^2/n}$$

and, finally, from (21):

$$| G_n | < e^{|v|^2\alpha^2 T^2} = e^{|u|\alpha^2 T^2}. \tag{26}$$

Thus it is seen that for

$$| u | < \frac{1}{2\alpha^2} \quad \text{or} \quad 1 - 2\alpha^2 | u | \geqq \eta^2 > 0 \tag{27}$$

the integral (15) admits the upper bound

$$| Q_n(u) | < (2\pi)^{-k/2} \iint \cdots \int e^{-\eta^2 T^2/2} \, dt_1, dt_2, dt_k = \eta^{-k}. \tag{28}$$

It also follows that the contribution to $Q_n(u)$ from the region $T > T_0$ tends to zero with increasing T_0, uniformly with respect to n and with respect to u in the region $| u | < 1/2\alpha^2$.

3. Asymptotic value of $Q_n(u)$. If the quantity F_ν introduced in (22) is considered as a function of $z_1/\sqrt{n}$, $z_2/\sqrt{n}$, $\cdots$, $z_k/\sqrt{n}$, we may write

$$F_\nu(z_1, z_2, \cdots, z_k) = \sum_{\kappa=1}^{k+1} p_{\nu\kappa} e^{v(z_\kappa - \bar{z}_\nu)}. \tag{29}$$

Here, $\bar{z}_\nu$ is defined by (20) and, on the right-hand side, z_{k+1} is zero. These functions $F_\nu(z_1, z_2, \cdots, z_k)$ for $\nu = 1, 2, 3, \cdots$ have all the properties required in Lemma C of Part I: At the point $z_1 = z_2 = \cdots = z_k = 0$ one has $F_\nu = 1$, the first derivatives are

$$\frac{\partial F_\nu}{\partial z_\iota} = vp_{\nu\iota} - vp_{\nu\iota} \sum_{\kappa=1}^{k+1} p_{\nu\kappa} = 0$$

and the second derivatives, ($\iota \neq \kappa$),

$$\begin{aligned} \frac{\partial^2 F_\nu}{\partial z_\iota^2} &= v^2 p_{\nu\iota}(1 - p_{\nu\iota}) - v^2 p_{\nu\iota} \left[p_{\nu\iota} - p_{\nu\iota} \sum_{\kappa=1}^{k+1} p_{\nu\kappa} \right] = v^2 p_{\nu\iota}(1 - p_{\nu\iota}) \\ \frac{\partial^2 F_\nu}{\partial z_\iota \partial z_\kappa} &= v^2 p_{\nu\iota}(-p_{\nu\kappa}) - v^2 p_{\nu\iota} \left[p_{\nu\kappa} - p_{\nu\kappa} \sum_{\lambda=1}^{k+1} p_{\nu\lambda} \right] = -v^2 p_{\nu\iota} p_{\nu\kappa}. \end{aligned} \tag{30}$$

The third derivatives are certainly bounded in any finite region of the z-space, and this means also in any finite region of the t-space.

The matrix of the second derivatives except for the factor v^2 is exactly that defined in eq. (19) of Part I:

$$P_{\iota\kappa}^{(\nu)} = p_{\nu\iota}\delta_{\iota\kappa} - p_{\nu\iota}p_{\nu\kappa} \tag{31}$$

and the arithmetical means of the derivatives from the matrix in eq. (23) of Part I:

$$\overline{P}_{\iota\kappa} = \frac{1}{n}\sum_{\nu=1}^{n} p_{\nu\iota}\,\delta_{\iota\kappa} - \frac{1}{n}\sum_{\nu=1}^{n} p_{\nu\iota}\,p_{\nu\kappa}. \tag{31'}$$

Applying Lemma C we find

$$G_n = G_n\left(\frac{z_1}{\sqrt{n}}, \frac{z_2}{\sqrt{n}}, \cdots, \frac{z_k}{\sqrt{n}}\right) \sim \exp\left[\frac{v^2}{2}\sum_{\iota,\kappa}\overline{P}_{\iota\kappa}z_\iota z_\kappa\right]. \tag{32}$$

This is valid in any finite t-region. Since it has been shown at the end of the foregoing section that, for small $|v|$, the outside contribution to the integral (15) converges uniformly (for all n) towards zero, we are allowed to introduce (32) in (15). Writing

$$\sum_{\iota,\kappa}\overline{P}_{\iota\kappa}z_\iota z_\kappa = \sum_{\iota,\kappa}\gamma_{\iota\kappa}t_\iota t_\kappa, \qquad \text{whereby } \gamma_{\iota\kappa} = \sum_{\lambda,\mu}\overline{P}_{\lambda\mu}\alpha_{\iota\lambda}\alpha_{\kappa\mu} \tag{33}$$

equation (15) becomes

$$Q_n(u) \sim (2\pi)^{-k/2}\iint\cdots\int\exp\left[-\tfrac{1}{2}\sum_{\kappa}t_\kappa^2 + \tfrac{1}{2}ui\sum_{\iota,\kappa}\gamma_{\iota\kappa}t_\iota t_\kappa\right]dt_1\,dt_2\cdots dt_k. \tag{34}$$

Now, it is well known that if $m_{\iota\kappa}$ is any positive definite matrix with the determinant $|m_{\iota\kappa}|$, then

$$(2\pi)^{-k/2}\iint\cdots\int\exp\left[-\tfrac{1}{2}\sum_{\iota,\kappa}m_{\iota\kappa}t_\iota t_\kappa\right]dt_1\,dt_2\cdots dt_k = \frac{1}{\sqrt{|m_{\iota\kappa}|}}. \tag{35}$$

This is likewise true if the matrix $m_{\iota\kappa}$, which we also call M, has the form $M = M_1 - \lambda M_2$ where M_1 is positive definite, M_2 arbitrary (complex) and $|\lambda|$ sufficiently small. Thus, the integration formula (35) applies to (34) and the result is reached, for small $|u|$:

$$Q_n(u) \sim Q(u) = \frac{1}{\sqrt{D(ui)}} \quad \text{with} \quad D(\lambda) = |\delta_{\iota\kappa} - \lambda\gamma_{\iota\kappa}|. \tag{36}$$

If the $\alpha_{\iota\kappa}$ which transform the given quadric into a sum of squares are known, (36) with (33) supply the solution of our problem.

The formula (36) is susceptible of several useful transformations. Let us write A for the matrix $\alpha_{\iota\kappa}$, A' for the transposed matrix, and Ψ, $\overline{P}$, Γ, I respectively for the matrices $\psi_{\iota\kappa}$, $\overline{P}_{\iota\kappa}$, $\gamma_{\iota\kappa}$, $\delta_{\iota\kappa}$. Then, obviously

$$\Psi = A'A, \qquad \Gamma = A\overline{P}A', \qquad M = I - ui\Gamma. \tag{37}$$

If we multiply M by A' to the left and by A to the right, we obtain

$$A'MA = A'IA - ui\,A'A\overline{P}A'A = \Psi - ui\,\Psi\overline{P}\Psi. \tag{38}$$

In this operation the determinant of M is multiplied by $|\psi_{\iota\kappa}|$. Thus $D(\lambda)$ can be written as

(39) $$D(\lambda) = \frac{|\psi_{\iota\kappa} - \lambda\gamma'_{\iota\kappa}|}{|\psi_{\iota\kappa}|} \quad \text{with} \quad \gamma'_{\iota\kappa} = \sum_{\lambda,\mu} \psi_{\iota\lambda} \bar{P}_{\lambda\mu} \psi_{\mu\kappa} .$$

Here, the knowledge of the $\alpha_{\iota\kappa}$ is no longer required.

If the matrix (38) is multiplied twice by Ψ^*, the inverse of Ψ, we find $\Psi^* - ui\bar{P}$ and, therefore,

(40) $$D(\lambda) = |\psi_{\iota\kappa}| \times |\psi^*_{\iota\kappa} - \lambda\bar{P}_{\iota\kappa}|.$$

As $\bar{P}$ is positive definite and Ψ^* real, it follows that all roots of $D(\lambda)$—the "Eigenwerte" of Γ—are real numbers. Therefore, $D^{-1/2}(ui)$ is a regular function along the real axis in the u-plane. Thus, (36) which was proved so far for small $|u|$ only remains valid for all real values of u: The c.f. of the asymptotic distribution is represented by $D^{-1/2}(ui)$ for all real u-values.

Multiplying (38) only once by Ψ^* we obtain one of the two forms

(41) $$I - ui\, \Psi\bar{P} \qquad \text{or} \qquad I - ui\, \bar{P}\Psi$$

which lead to

(42) $$D(\lambda) = |\delta_{\iota\kappa} - \lambda s_{\iota\kappa}| = |\delta_{\iota\kappa} - \lambda s_{\kappa\iota}|, \qquad s_{\iota\kappa} = \sum_{\mu} \psi_{\iota\mu} \bar{P}_{\mu\kappa} .$$

Although this formula has been derived by means of Ψ^* it can be seen by continuity considerations that it remains valid whatever the (symmetric) matrix $\psi_{\iota\kappa}$ is. The formula makes it clear that the asymptotic distribution of the quadric $\Sigma\psi_{\iota\kappa}\xi_\iota\xi_\kappa$ is completely determined by the "Eigenwerte" of the matrix $S = \Psi\bar{P}$. This bears out our second main theorem in Chapter II, as far as quartics of the form (8) are concerned. It will be seen in sec. 5 how (42) applies to the continuous case.

We, finally, apply to (36) a transformation that is valid only if $\bar{P}$ has an inverse matrix $\bar{P}^*$. (As shown in Part I, sec. 4 this is not the case if the k intervals to which the subscripts $1, 2, \cdots, k$ refer cover the whole range of the variables $x_1, x_2, \cdots, x_n$). Multiplying (41) by $\bar{P}^*$ we find the matrix $\bar{P}^* - ui\Psi$ and thus

(43) $$D(\lambda) = |\bar{P}_{\iota\kappa}| \times |\bar{P}^*_{ik} - \lambda\psi_{\iota\kappa}|.$$

This is equivalent to

(44) $$Q(u) = |\bar{P}_{\iota\kappa}|^{1/2} \int\int \cdots \int \exp [- \tfrac{1}{2}\Sigma\bar{P}^*_{\iota\kappa}\xi_\iota\xi_\kappa + \tfrac{1}{2}ui\sum\psi_{\iota\kappa}\xi_\iota\xi_\kappa]\, d\xi_1 d\xi_2 \cdots d\xi_k .$$

According to the definition of the characteristic function eq. (44) can be interpreted as stating that

(45) $$|\bar{P}_{\iota\kappa}|^{\frac{1}{2}} \exp [-\tfrac{1}{2}\Sigma\bar{P}^*_{\iota\kappa}\, \xi_\iota\xi_\kappa]$$

is the asymptotic probability density for the simultaneous occurrence of ξ_1, $\xi_2, \cdots, \xi_k$. The expression (45) can be arrived at by applying the Central Limit Theorem to the case of k independent chance variables. Since, however, $\bar{P}^*$ does not exist in general, eq. (44) would not be a suitable point of departure for developing the theory that concerns us here.

4. Asymptotic value of $P_n(x)$, illustrations. The relationship between the c.f. and the c.d.f. of a distribution is well known and need not be discussed here in detail. We shall use, in this section, two aspects of this relationship only. First, the continuity theorem, first proved by G. Pólya [5], stating that if the c.f. $Q_n(u)$ tend towards a limiting function $Q(u)$, the corresponding c.d.f. $P_n(x)$ tend towards the $P(x)$ that corresponds to $Q(u)$. Second, the additivity, i.e. if $Q(u)$ is of the form $\alpha Q'(u) + \beta Q''(u)$ with $\alpha + \beta = 1$, then $P(x)$ is $\alpha P'(x) + \beta P''(x)$ with the $P'(x)$, $P''(x)$ corresponding to $Q'(u)$ and $Q''(u)$ respectively. The following three groups of examples will illustrate the application of the foregoing results.

a) Let us first consider a function of two excess values ξ_1, ξ_2 only

$$x = \frac{n}{2} f = \frac{n}{2} (A\xi_1^2 + 2B\xi_1\xi_2 + C\xi_2^2) \tag{46}$$

where the matrix Ψ is given by $\Psi_{11} = A$, $\Psi_{12} = \Psi_{21} = B$, $\Psi_{22} = C$. The product matrix $\bar{P}\Psi$ is

$$\begin{matrix} A\bar{P}_{11} + B\bar{P}_{12} & B\bar{P}_{11} + C\bar{P}_{12} \\ A\bar{P}_{21} + B\bar{P}_{22} & B\bar{P}_{21} + C\bar{P}_{22} \end{matrix} \tag{47}$$

and the determinant of $I - \lambda\bar{P}\Psi$

$$D(\lambda) = 1 - \lambda[A\bar{P}_{11} + 2B\bar{P}_{12} + C\bar{P}_{22}] + \lambda^2(AC - B^2)(\bar{P}_{11}\bar{P}_{22} - \bar{P}_{12}^2). \tag{48}$$

If λ_1, λ_2 are the two real roots of $D(\lambda) = 0$, the asymptotic probability density of x will be

$$\frac{dP(x)}{dx} = \frac{1}{2\pi} \int \frac{e^{-uix}\, du}{\sqrt{\left(1 - \frac{ui}{\lambda_1}\right)\left(1 - \frac{ui}{\lambda_2}\right)}}. \tag{49}$$

We are particularly interested in the case that $\bar{P}$ is "complete," i.e. a matrix with all horizontal and vertical sums vanishing. Then $\bar{P}_{11} = \bar{P}_{12} = \bar{P}_{22} = \overline{p_1p_2}$, the last term in (48) cancels out and the only Eigenwert is $\lambda_1 = 1/(A - 2B + C)\overline{p_1p_2}$. Here, instead of (49) we have

$$\frac{dP(x)}{dx} = \frac{1}{2\pi} \int \frac{e^{-uix}\, du}{\sqrt{1 - \frac{ui}{\lambda_1}}} = \sqrt{\frac{\lambda_1}{\pi}} \frac{e^{-\lambda_1 x}}{\sqrt{x}} \tag{50}$$

This is, with respect to $\sqrt{|x|}$ a Gauss distribution with the variance $|A - 2B + C| \overline{p_1p_2}/2$.

If, in addition to the assumption that $\bar{P}$ is "complete" (i.e. in the present case that $p_{\nu 1} + p_{\nu 2} = 1$ for all ν) the further assumption is made that the two intervals I_1 and I_2 cover the whole range of the original chance variables $x_1, x_2, x_3, \cdots$, one would have also $\xi_1 + \xi_2 = 0$ and from (46)

$$x = \frac{n}{2}(A - 2B + C)\xi_1^2.$$

In this case, $\sqrt{|x|}$ is a linear statistical function and the Central Limit Theorem leads to the same result as that expressed in (50). It is seen, however, from our derivation, that (50) holds under wider conditions: If $p_{\nu 1} + p_{\nu 2} = 1$ for all ν, there may exist another interval I_3 within the range of the chance variables $x_1, x_2, x_3, \cdots$ so that $\xi_1 + \xi_2$ is not necessarily zero.

The latter remark suggests the following general theorem: If f is a function of the k variables $\xi_1, \xi_2, \cdots, \xi_k$ and g another such function but vanishing when $\xi_1 + \xi_2 + \cdots + \xi_k = 0$, then f and $f + g$ have the same asymptotic distribution provided that for each ν the sum $p_{\nu 1} + p_{\nu 2} + \cdots + p_{\nu k} = 1$. In the case of quadrics this result is equivalent to the following matrix theorem: If $\bar{P}$, Ψ, A are symmetric matrices, $\bar{P}$ with all horizontal and vertical sums equal to zero, Ψ arbitrary, and A of the form $a_{\iota\kappa} = a_\iota + a_\kappa$ then the two products

$$\bar{P}\Psi \quad \text{and} \quad \bar{P}(\Psi + A) \tag{51}$$

have the same characteristic roots.—This can be proved by the usual methods of matrix calculus. The matrix $\bar{P}A$ has all characteristic roots equal to zero.[2]

b) In the definition of Karl Pearson's test function which is usually called χ^2, it is presumed that a sample is drawn from the combination of n *equal* distributions. In this case all $P^{(\nu)}$ are equal and coincide with $\bar{P}$ which then can simply be written P:

$$P_{\iota\kappa} = p_\iota \delta_{\iota\kappa} - p_\iota p_\kappa. \tag{52}$$

The chance variable we now consider will be

$$x = \frac{n}{2} f = \frac{n}{2} \sum_\iota \frac{\xi_\iota^2}{p_\iota} = \tfrac{1}{2}\chi^2. \tag{53}$$

Thus $\psi_{\iota\kappa} = \delta_{\iota\kappa}/p_\iota$ and the elements of $P\Psi$ are

$$(P\Psi)_{\iota\kappa} = \sum_\mu P_{\iota\mu}\psi_{\mu\kappa} = \delta_{\iota\kappa} - p_\iota. \tag{53'}$$

The matrix $I - \lambda P\Psi$ has the elements

$$\delta_{\iota\kappa}(1 - \lambda) + \lambda p_\iota.$$

If the kth column is subtracted from any one of the others, only two terms remain, one equal to $1 - \lambda$ and one equal $-(1 - \lambda)$ in the last row. Thus, the

[2] A proof of the matrix theorem has meanwhile been published by Alfred Brauer, *Bull. Amer. Math. Soc.*, Vol. 53 (1947), pp. 605–607.

determinant $D(\lambda)$ includes $(k - 1)$ times the factor $(1 - \lambda)$. On the other hand, $D(\lambda)$ is of degree $(k - 1)$ and has the absolute term 1. Therefore

$$D(\lambda) = (1 - \lambda)^{k-1}. \tag{54}$$

This supplies the χ^2-distribution with $(k - 1)$ "degrees of freedom"

$$Q(u) = (1 - ui)^{-\frac{k-1}{2}}, \quad \frac{dP(x)}{dx} = \frac{1}{\Gamma\left(\frac{k-1}{2}\right)} x^{\frac{k-3}{2}} e^{-x}, \qquad (x \geqq 0). \tag{55}$$

Again, our result is slightly more general than that reached in the usual theory. It includes the case that in addition to the k intervals with the probabilities $p_1, p_2, \cdots, p_k$ (whose sum is 1) there are other intervals with probability zero. On the other hand, if to χ^2 a term of the form $n\Sigma(a_\iota + a_\kappa)\xi_\iota\xi_\kappa$ is added, this would not change the asymptotic distribution.

One may ask for other quadratic functions of $\xi_1, \xi_2, \cdots, \xi_k$ whose asymptotic distribution is given by (55). In particular, one might be interested in a generalization of χ^2 for the case of *unequal original distributions*. The answer can easily be given by introducing the cofactors of order $(k - 1)$ and of order $(k - 2)$ of the determinant $|\bar{P}_{\iota\kappa}|$. It was mentioned in sec. 4 of Part I that all cofactors of order $(k - 1)$—in the case of "complete" $\bar{P}$—have the same value. It may be denoted by Δ. The cofactor corresponding to the lines ι, κ and the columns λ, μ will be denoted by $\Pi_{\iota\kappa;\lambda\mu}$ with $\Pi = 0$ if $\iota = \kappa$ or $\lambda = \mu$. Then, if l is any one of the integrers $1, 2, \cdots, k$

$$\psi_{\iota\kappa} = \frac{1}{\Delta} \Pi_{\iota l;\kappa l}; \qquad \iota, \kappa \neq l \tag{56}$$

is one possible solution. In fact, the product $\bar{P}\Psi$ has in this case the elements

$$\begin{aligned} (\bar{P}\Psi)_{\iota\kappa} &= \delta_{\iota\kappa}, \quad \text{for} \quad \iota, \kappa \neq l \\ &= -1, \quad \text{"} \quad \iota = l, \kappa \neq l \\ &= 0, \quad \text{"} \quad \kappa = l \end{aligned} \tag{57}$$

The determinant of $I - \lambda\bar{P}\Psi$ is then seen to equal $(1 - \lambda)^{k-1}$.

The solution (56), however, is unsymmetrical in the sense that it does not include any terms with ξ_l. A completely symmetrical solution in which all ξ play the same role is given by

$$\psi_{\iota\kappa} = \frac{1}{k\Delta} \sum_{l=1}^{k} \Pi_{\iota l;\kappa l} \tag{58}$$

According to (57) the matrix $\bar{P}\Psi$ now consists of terms $(k - 1)/k$ in the principal diagonal and $-1/k$ at all other places, that is

$$(P\Psi)_{\iota\kappa} = \delta_{\iota\kappa} - \frac{1}{k}. \tag{58'}$$

In the same way as in the case of (53′) it can be seen that the determinant of $I - \lambda \bar{P}\Psi$ equals here $(1 - \lambda)^{k-1}$. *The asymptotic distribution of* $\sum \psi_{\iota\kappa}\xi_\iota\xi_\kappa$ *with the coefficients* (58) *is, therefore, the* χ^2*-distribution with* $(k - 1)$ *degrees of freedom.*

If the formula (58) is applied to the case of equal $P^{(\nu)}$ the corresponding quadric becomes

$$\sum_\iota \frac{1}{p_\iota} \xi_\iota^2 + \frac{1}{k} \sum_\iota \frac{1}{p_\iota} \left(\sum_\iota \xi_\kappa\right)^2,$$

that is, χ^2 + a term vanishing with $\xi_1 + \xi_2 + \cdots + \xi_k$. One can easily modify (58) so that it leads to χ^2 without any addition.

c) A third group of examples where the asymptotic density is expressed by simple functions is that where $D(\lambda)$ is an exact square, that is, all characteristic roots (except the one that is zero) have even multiplicities. Let us assume $k = 2m + 1$ and let $\lambda_1, \lambda_2, \cdots, \lambda_m$ be m double roots. Then

$$Q(u) = \prod_{\mu=1}^{m} \left(1 - \frac{ui}{\lambda_\mu}\right)^{-1} = \sum_{\mu=1}^{m} \frac{\lambda_\mu A_\mu}{\lambda_\mu - ui} \tag{59}$$

with

$$A_\mu = \prod_{\iota \neq \mu}^{1 \cdots m} \left(1 - \frac{\lambda_\mu}{\lambda_\iota}\right) \tag{59}$$

and therefore

$$\frac{dP(x)}{dx} = \sum_{\mu=1}^{m} A_\mu \lambda_\mu e^{-\lambda_\mu x}, \qquad x \geqq 0. \tag{60}$$

Assume, for instance, that all original distributions are uniform, that is

$$P_{\iota\kappa}^{(\nu)} = P_{\iota\kappa} = \frac{1}{k} \delta_{\iota\kappa} - \frac{1}{k^2}$$

and that the quadric f is given in the form (11) with the following $\alpha_{\iota\kappa}$:

$$\begin{aligned} \alpha_{\iota\kappa} &= \sqrt{kc_1} && \text{for } \iota = 1 \\ &= \sqrt{kc_\iota} && \text{"} \ \iota > 1, \kappa = 1, 2, \cdots, \iota - 1 \\ &= -(\iota - 1)\sqrt{kc_\iota} && \text{"} \ \iota > 1, \kappa = \iota \\ &= 0 && \text{"} \ \iota > 1, \kappa = \iota + 1, \iota + 2, \cdots, k. \end{aligned} \tag{61}$$

Then, the $\gamma_{\iota\kappa}$ as defined in (33) become

$$\begin{aligned} \gamma_{\iota\kappa} &= c_\iota \iota(\iota - 1)\delta_{\iota\kappa} && \text{for } \iota \text{ or } \kappa > 1 \\ &= 0 && \text{"} \ \iota = \kappa = 1 \end{aligned} \tag{62}$$

and $D(\lambda)$ according to (36) takes the form

$$D(\lambda) = | \delta_{\iota\kappa} - \lambda\gamma_{\iota\kappa} | = \prod_{\iota=2}^{k} [1 - \lambda c_\iota \iota(\iota - 1)]. \tag{63}$$

In other terms, for the quadric

$$f = kc_1(\xi_1 + \cdots + \xi_k)^2 + kc_2(\xi_1 - \xi_2)^2 + kc_3(\xi_1 + \xi_2 - 2\xi_3)^2 + \cdots + kc_k[\xi_1 + \xi_2 + \cdots + \xi_{k-1} - (k - 1)\xi_k]^2$$

the characteristic λ-values are $1/c_\iota\iota(\iota - 1)$.

Now, to obtain the case of m double roots with $k = 2m + 1$ we have simply to choose

$$c_2 = 3c_3\,,\ 3c_4 = 5c_5\,,\ 5c_6 = 7c_7\,,\ \cdots .$$

The first term on the right-hand side can be entirely omitted in accordance to what was said in connection with (51). Besides, for the same reason, the expression can be simplified in various ways by assuming $\xi_1 + \xi_2 + \cdots + \xi_k = 0$.

As a numerical example, take $k = 5$, $c_2 = 3$, $c_3 = 1$, $c \ = 5$, $c_5 = 3$. Then

$$f = 20(\xi_1^2 + \xi_2^2 + \xi_3^2 + 20\,\xi_4^2 + 20\,\xi_5^2 - \xi_1\xi_2 - \xi_2\xi_3 - \xi_3\xi_1 + 10\,\xi_4\xi_5)$$

leads to the characteristic values $\lambda = 1/6$ and $1/60$ and the asymptotic density becomes

$$\frac{dP}{dx} = \frac{1}{54}\left(e^{-x/60} - e^{-x/6}\right).$$

In a similar way other groups of quadrics with asymptotic distributions of the type (60) can easily be constructed. One may, for instance, use eq. (41) and make vanish, in the matrix $S = \bar{P}\Psi$, all elements on one side of the diagonal so that the roots are immediately known.

5. Transition to the continuous case. In this concluding section, the transition to the case of a quadric of the form (1) with continuous $\psi\ (x, y)$ will be outlined. The formula best fit for this purpose is eq. (36). We therefore suppose the statistical function f given as

$$(64)\quad f = \iint \psi(x, y)\, dT_n(x)\, dT_n(y) \quad \text{with} \quad \psi(x, y) = \int \alpha(r, x)\alpha(r, y)\, dr.$$

In analogy to (33) we derive

$$(65)\quad \begin{aligned}\gamma(x, y) &= \iint \alpha(x, s)\alpha(y, t)\, d\bar{U}_n(s, t) \\ &= \int \alpha(x, s)\alpha(y, s)\, d\bar{V}_n(s) - \iint \alpha(x, s)\alpha(y, t)\, d\overline{W}_n(s, t).\end{aligned}$$

Since $d\overline{W}$ is symmetric, this function $\gamma(x, y)$ is symmetric with respect to x and y. If $D(\lambda)$ denotes the *Fredholm determinant* of the "kernel" $\gamma(x, y)$, we conclude from (36) that the characteristic function of the asymptotic distribution of f will be given by

$$Q_n(u) \sim \frac{1}{D(ui)} \tag{66}$$

if certain convergence conditions are satisfied.

In order to establish (66) the main point is to find a sequence of functions $\psi_1(x, y), \psi_2(x, y), \cdots$ each of the type considered in the foregoing Sections and such that 1) the distribution of the quadric f_k with the coefficients ψ_k tends towards the distribution of f with increasing k and independently of n; and 2) that the determinants D_k corresponding to ψ_k converge towards D as k increases indefinitely. Using our Lemma A we can replace the first condition by asking that the expectation of $|f - f_k|$ should go to zero with $k \to \infty$ independently of n.

The following assumptions shall be made concerning f and the $V_\nu(x)$: The function $\alpha(r, x)$ in (64) is continuous and bounded in every finite region; there exist two positive continuous functions $\alpha(r)$, $\beta(x)$ such that

$$|\alpha(r, x)| \leqq \alpha(r)\beta(x) \tag{67}$$

and that the integrals

$$\int \alpha^2(r)\, dr = M, \qquad \int \beta(x)\, dV_\nu(x), \qquad \int \beta^2(x)\, dV_\nu(x) \tag{68}$$

exist, the latter two being bounded and converging uniformly with respect to ν. We are going to devise a step function $\psi_k(x, y)$ so that for the corresponding f_k and any positive ϵ_1

$$E\{\, |f - f_k|\, \} \leqq \epsilon_1 . \tag{69}$$

Let N be an upper bound of the integrals

$$\int \beta(x)\, dV_\nu(x) \leqq N, \qquad \int \beta(x)\, d\bar{V}_n(x) \leqq N \tag{70}$$

and $\epsilon = \epsilon_1/(5 + 8\, N)$. Choose a value L such that

$$\int_{|x|>L} \beta(x)\, dV_\nu(x) \leqq \frac{\epsilon}{M}, \qquad \int_{|x|>L} \beta^2(x)\, dV_\nu(x) \leqq \frac{\epsilon}{M} \tag{71}$$

and, calling B the maximum of $\beta(x)$ in $|x| \leqq L$, another quantity R such that

$$\int_{|r|>R} \alpha^2(r)\, dr \leqq \frac{\epsilon}{2B^2}. \tag{72}$$

We subdivide, in the x-y-r-space, the domain $|x| \leqq L, |y| \leqq L, |r| \leqq R$ in k^3 equal cells where k is determined by the condition that the absolute value of the variation of $\alpha(r, x)\alpha(r, y)$ within each cell does not exceed $\epsilon/4R$. Outside this domain we set $\psi_k(r, x) = 0$ while inside the domain $\alpha_k(r, x)\alpha_k(r, y)$ shall equal the value that $\alpha(r, x)\alpha(r, y)$ assumes in the center of the respective cell. Then $\psi_k(x, y)$ will be defined by

$$\psi_k(x, y) = \int \alpha_k(r, x)\alpha_k(r, y)\, dr. \tag{73}$$

From the definition of k and from (67) and (72) it follows that

$$\begin{aligned} |\psi(x, y) - \psi_k(x, y)| &\leqq \int_{|r| \leqq R} |\alpha(r, x)\alpha(r, y) - \alpha_k(r, x)\alpha_k(r, y)|\, dr \\ &\quad + \int_{|r|>R} |\alpha(r, x)\alpha(r, y)|\, dr \\ &\leqq 2R\frac{\epsilon}{4R} + \beta(x)\beta(y)\int_{|r|>R} \alpha^2(r)\, dr \leqq \frac{\epsilon}{2} + B^2\frac{\epsilon}{2B^2} = \epsilon \end{aligned} \tag{74}$$

as long as $|x| \leqq L, |y| \leqq L$. If this square is called (L) and the complementary region $(\bar{L})$ we have

$$\begin{aligned} f - f_k = \iint_{(L)} [\psi(x, y) - \psi_k(x, y)]\, dT_n(x)\, dT_n(y) \\ + \iint_{(\bar{L})} \psi(x, y)\, dT_n(x)\, dT_n(y) \end{aligned} \tag{75}$$

and since the integral of $|dT_n(x)\, dT_n(y)|$ is not larger than 4, while, according to (64) and (67)

$$|\psi(x, y)| \leqq \beta(x)\beta(y) \int \alpha^2(r)\, dr = M\beta(x)\beta(y) \tag{76}$$

we conclude from (74) and (75)

$$|f - f_k| \leqq 4\epsilon + M \iint_{(\bar{L})} \beta(x)\beta(y)\, |dT_n(x)\, dT_n(y)|. \tag{77}$$

This gives

$$E\{|f - f_k|\} \leqq 4\epsilon + M \iint_{(\bar{L})} \beta(x)\beta(y) E\{|dT_n(x)\, dT_n(y)|\}. \tag{78}$$

Now, from $|dT_n| = |dS_n - d\bar{V}_n| \leqq dT_n + 2d\bar{V}_n$ and from the formulas derived in Part II,

$$E\{dT_n(x)\} = 0, \qquad E\{dT_n(x)\, dT_n(y)\} = \frac{1}{n}\, d\bar{U}_n(x, y)$$

it follows

$$E\{|dT_n(x)\, dT_n(y)|\} \leqq \frac{1}{n}\, d\bar{U}_n(x, y) + 4\, d\bar{V}_n(x)\, d\bar{V}_n(y) \tag{79}$$

with

$$d\bar{U}_n(x, y) = \delta(x, y)\, d\bar{V}_n(x) - d\overline{W}_n(x, y) \leqq \delta(x, y)\, d\bar{V}_n(x). \tag{79'}$$

If this is introduced in (78) and (71) taken into account, we find

$$\begin{aligned} E\{\,|f - f_k|\,\} &\leqq 4\epsilon + M\,\frac{1}{n}\int_{|x|>L} \beta^2(x)\, d\bar{V}_n(x) \\ &\qquad + 4M \iint_{(\bar{L})} \beta(x)\beta(y)\, d\bar{V}_n(x)\, d\bar{V}_n(y) \\ &\leqq 4\epsilon + \frac{1}{n}\epsilon + 4 \times 2N\epsilon \leqq (5 + 8N)\epsilon = \epsilon_1 \end{aligned} \tag{80}$$

as required in (69).

On the other hand, it can be seen that the kernel $\gamma(x, y)$ as defined in (65) is the limit of the sequence $\gamma_k(x, y)$

$$\begin{aligned} \gamma_k(x, y) &= \iint_{(L)} \alpha_k(x, s)\alpha_k(y, t)\, d\bar{U}_n(s, t) && \text{for } x, y \text{ in } (R) \\ &= 0 && \text{for } x, y \text{ in } (\bar{R}) \end{aligned} \tag{81}$$

where (R) means the region $|\,x\,| \leqq R$, $|\,y\,| \leqq R$ and $(\bar{R})$ the complementary region. In fact, from the definition of k and eqs. (67) and (71) one has for x, y in (R):

$$\begin{aligned} |\,\gamma(x, y) - \gamma_k(x, y)\,| &\leqq \frac{\epsilon}{4R}\iint_{(L)} |\,d\bar{U}_n(s, t)\,| \\ &\qquad + \iint_{(\bar{L})} |\,\alpha(x, s)\alpha(y, t)\, d\bar{U}_n(s, t)\,| \\ &\leqq \frac{\epsilon}{2R} + \alpha(x)\alpha(y)\left[\int_{|s|>L} \beta^2(s)\, d\bar{V}_n(s) \right. \\ &\qquad \left. + \frac{1}{n}\sum_{\nu=1}^{n} \iint_{(\bar{L})} \beta(s)\beta(t)\, dV_\nu(s)\, dV_\nu(t) \right] \\ &\leqq \frac{\epsilon}{2R} + \alpha(x)\alpha(y)\,\frac{\epsilon}{M}\,(1 + 2N). \end{aligned} \tag{82}$$

Since $\alpha(x)$ is bounded, the right-hand side goes to zero with ϵ. Finally, for x, y in $(\bar{R})$ we have

$$\begin{aligned} |\,\gamma(x, y) - \gamma_k(x, y)\,| &\leqq \iint |\,\alpha(x, s)\alpha(y, t)\, d\bar{U}_n(s, t)\,| \\ &\leqq \alpha(x)\alpha(y)\left[\int \beta^2(s)\, d\bar{V}_n(s) \right. \\ &\qquad \left. + \frac{1}{n}\sum_{\nu=1}^{n} \iint \beta(s)\beta(t)\, dV_\nu(s)\, dV_\nu(t).\right] \end{aligned} \tag{83}$$

Here, the two terms in the brackets are bounded, but $\alpha(x)\alpha(y)$ goes to zero as R increases. The conclusion is that $\gamma_k(x, y)$ tends uniformly towards $\gamma(x, y)$ with $k \to \infty$.

Thus, eq. (66) is established provided that the function $\gamma(x, y)$ defined in (65) has a Fredholm determinant $D(\lambda)$ that is the limit of the corresponding algebraic determinants and provided that the c.f. $\sqrt{1/D(ui)}$ leads to a c.d.f. with bounded derivative.

As an *example* let us consider the case

$$\begin{aligned} \alpha(r, x) &= \sqrt{g'(r)} \text{ for } r \geqq x \\ &= 0 \qquad \text{`` } r < x. \end{aligned} \tag{84}$$

This function is not continuous as it was assumed in establishing (66). However, the existence of a single discontinuity line, $x = r$, does not invalidate the argument. We assume $g'(r) = 0$ and equal to dg/dr. Then, in the case of (84):

$$\begin{aligned} \psi(x, y) = \int \alpha(r, x)\alpha(r, y)dr &= -g(y) \quad \text{for } x \leqq y \\ &= -g(x) \quad \text{`` } x \geqq y. \end{aligned} \tag{85}$$

Since, however, adding to ψ a function of x or of y alone does not change the value of f, we can also use

$$\begin{aligned} \psi(x, y) &= g(x) \quad \text{for } x \leqq y \\ &= g(y) \quad \text{`` } x \geqq y. \end{aligned} \tag{85'}$$

The statistical function f that corresponds to (84) can be computed either from (85) or (85′)—or directly from (84) if we use the formula that follows from (64)

$$f = \int \left[\int \alpha(r, x)\, dT_n(x) \right]^2 dr. \tag{86}$$

The integral in the brackets is, in our case, seen to equal $\sqrt{g'(r)}\ T_n(r)$, thus

$$f = \int g'(r)[S_n(r) - \bar{V}_n(r)]^2\, dr. \tag{86'}$$

This is exactly the test function ω^2 mentioned in the Introduction, eq. (3).

To find the distribution of f we have to compute $\gamma(x, y)$. Its definition (65) can be written in the form

$$\gamma(x, y) = \frac{1}{n} \sum_{\nu=1}^{n} \left[\int \alpha(x, s)\alpha(y, s)\, dV_\nu(s) - \int \alpha(x, s)\, dV_\nu(s) \int \alpha(y, s)\, dV_\nu(s) \right]. \tag{87}$$

This supplies in the case of (84)

$$\begin{aligned} \gamma(x, y) &= \sqrt{g'(x)g'(y)}[\bar{V}_n(x) - \overline{V_n(x)V_n(y)}] \text{ for } x \leqq y \\ &= \sqrt{g'(x)g'(y)}[\bar{V}_n(y) - \overline{V_n(x)V_n(y)}] \text{ `` } x \geqq y. \end{aligned} \tag{88}$$

Here, the second term in the brackets is the arithmetical mean of the products $V_\nu(x)V_\nu(y)$.

If the distributions $V_\nu(x)$ are all equal (independent of ν) we have simply to write $V(x)$ instead of $\bar{V}_n(x)$ and $V(x)V(y)$ instead of $\overline{V_n(x)V_n(y)}$. If, in addition, the distribution in the original collectives are uniform in the basic interval 0 to 1, one has

$$\gamma(x, y) = \sqrt{g'(x)g'(y)}\, x(1 - y) \text{ for } 0 \leqq x \leqq y \leqq 1$$
$$(89) \qquad = \sqrt{g'(x)g'(y)}\, y(1 - x) \quad \text{"} \quad 0 \leqq y \leqq x \leqq 1.$$

This is the case dealt with in Smirnoff's papers [7, 8]. If, finally, $g'(x)$ is supposed to be equal to 1 in the interval 0, 1, we arrive at a kernel $\gamma(x, y)$ whose Fredholm determinant is well known:

$$(90) \qquad \begin{aligned} \gamma(x, y) &= x(1 - y) \quad \text{for} \quad x \leqq y \\ &= y(1 - x) \quad \text{"} \quad x \geqq y. \end{aligned} \qquad D(\lambda) = \frac{\sin\sqrt{\lambda}}{\sqrt{\lambda}}.$$

This supplies immediately the c.f. and (in form of a definite integral) the c.d.f. of the asymptotic distribution of ω^2 for $g' = 1$.

The same result can be reached without the use of $\alpha(r, x)$ if we apply one of the transformations discussed in the foregoing Section. Take, for instance, instead of $\gamma(x, y)$ the unsymmetric kernel $\sigma(x, y)$ corresponding to the matrix $S = \bar{P}\Psi$ defined in (41). If all original distributions are equal, the element of S can be written as

$$(91) \qquad s_{\iota\kappa} = \sum_\mu P_{\iota\mu}\psi_{\mu\kappa} = p_\iota(\psi_{\iota\kappa} - \sum_\mu \psi_{\mu\kappa}p_\mu).$$

Calling $v(x)$ the density $dV(x)/dx$ in the continuous case, the corresponding kernel becomes

$$(92) \qquad \sigma(x, y) = v(x)\left[\psi(x, y) - \int \psi(s, y)v(s)\, ds\right].$$

With the ψ-values from (85′), $g' = 1$, $v = 1$, this gives

$$(92') \qquad \begin{aligned} \sigma(x, y) &= x - y + \frac{y^2}{2} \text{ for } \quad x \leqq y \\ &= \frac{y^2}{2} \qquad \text{"} \quad x \geqq y. \end{aligned}$$

It can easily be seen that the "Eigenfunctions" of this $\sigma(x, y)$ are $\sin(\sqrt{\lambda_m}\, x)$ with $\lambda_m = m^2\pi^2$, and, therefore, the Fredholm determinant is that indicated in (90).

It might be added that the expectation and the asymptotic variance of ω^2 can be computed, independently of the distribution, from the formulas developed in Part I. The results are

$$(93) \qquad nE\{\omega^2\} = \int g'(x)\overline{V_n(x)[1 - V_n(x)]}\, dx$$

and, in the case of all $V_\nu(x)$ equal

$$n^2 \mathrm{Var}\{\omega^2\} \sim 4 \iint\limits_{x \leq y} g'(x) g'(y) V^2(x) [1 - V(y)]^2 \, dx \, dy. \tag{94}$$

These formulas have already been given in [4].

Another, more general, remark is this. If all $V_\nu(x)$ are equal, one can reduce the problem, by a transformation of the original chance variable x into $x' = V(x)$, to the case of a uniform distribution over the interval 0 to 1. If the $V_\nu(x)$ are not equal, it might still be possible to find a transformation $x' = x'(x)$ such that all original distributions extend over a finite region on the x'-axis only. In this case the restrictions concerning the behavior of the distributions at infinity drop out.

REFERENCES

[1] Harald Cramér, "On the composition of elementary errors," *Skand. Aktuarietidsdrift*, Vol. 11 (1928), pp. 13–74, 141–180.

[2] R. v. Mises, "Les lois de probabilité pour les fonctions statistiques," *Ann. de l'Inst. Henri Poincaré*, Vol. 6 (1936), pp. 185–212.

[3] ———, "Sur les fonctions statistiques," *Soc. math. de France*, Conférence de la Réunion internat. des Mathématiciens, Paris, 1937.

[4] ———, *Wahrscheinlichkeitsrechnung und ihre Anwendung*, Leipzig and Wien, 1931.

[5] G. Pólya, "Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung," *Math. Zeitschr.*, Vol. 8 (1920), pp. 171–181.

[6] T. A. Shohat and T. D. Tamarkin, *The Problem of Moments*, Math. Surveys No. 1, New York, 1943.

[7] N. V. Smirnoff, "On the distribution of the ω^2-criterion of Mises," (In Russian), *Recueil Math.*, nouvelle série, Vol. 2 (1937), pp. 973–993.

[8] ———, "Sur la distribution de ω^2 (criterium de M. von Mises)," *Comptes Rendus Paris*, Vol. 202 (1936), p. 449.

[9] Vito Volterra. *Leçons sur les Fonctions de Ligne*, Paris, 1913.

[10] Vito Volterra and Joseph Pérès, *Théorie générale des Fonctionelles*, Paris, 1936.