On the algebra of distributions.

By ALFRÉD RÉNYI in Debrecen.

Introduction.

A real function $F(x)$ defined on the whole real axis which is non-decreasing, continuous to the right and which satisfies the conditions $F(-\infty)=0$, $F(+\infty)=1$ is called a *distribution function* (abbreviated d. f.). If $F_1(x)$ and $F_2(x)$ are both d. f.-s, we define their "*composition*" $F_1(x)*F_2(x)$ by

$$F_1(x)*F_2(x)=\int_{-\infty}^{+\infty}F_1(x-y)\,dF_2(y). \tag{1}$$

Clearly $F(x)=F_1(x)*F_2(x)$ defined by (1) is also a d. f. This operation of composition is evidently commutative and associative; thus the family of all d. f.-s form a semi-group with respect to composition defined by (1). This semi-group has a unity element $E(x)$ defined by $E(x)=0$ for $x<0$, $E(x)=1$ for $x\geqq 0$. By the algebra of distributions we mean the algebra of this semi-group.[1]) The algebra of distributions is of central importance in probability theory, as the composition of distributions is closely connected with the addition of independent random variables. If ξ_1 and ξ_2 are independent random variables having the d. f.-s $F_1(x)$ resp. $F_2(x)$ the d. f. of $\xi_1+\xi_2$ is given by (1). In probability theory usually the distributions of the variables $\xi_1,\xi_2,\ldots,\xi_n$ are known and it is asked for the distribution of the sum $\sum_{k=1}^{n}\xi_k$ or after some properties of this distribution. As a matter of fact the various well known forms of the laws of large numbers and of the central limit theorem are answers to questions of this type. Besides these problems of *composition*, problems of a different kind, problems of *factorization* have also been considered. By a factorization problem we understand a problem, in which the distribution of a sum of random variables is known, and — under some additional conditions — the distributions of the components are to be deter-

[1]) Cf. E. HILLE: Functional analysis and semi-groups, *American Math. Soc. Colloquium Publications*, Vol. 31 (1948).

mined. As an example we mention the following remarkable theorem of H. CRAMÉR[2]): Every factor of a normal distribution is also normal; by other words if $F(x)=F_1(x)*F_2(x)$ and $F(x)$ is the normal d. f. with parameters $(m,\ \sigma)$ i. e. if

$$F(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x} e^{-\frac{(y-m)^2}{2\sigma^2}}\,dy \tag{3}$$

it follows that $F_1(x)$ and $F_2(x)$ are also normal distributions with parameters $(m_1,\ \sigma_1)$ and $(m_2,\ \sigma_2)$, (clearly we have $m_1+m_2=m$ and $\sigma_1^2+\sigma_2^2=\sigma^2$). A similar result has been proved by D. RAIKOV[3]) for Poisson distributions. An other question which has been thoroughly investigated, concerns the factorization of distributions into an arbitrary great number of equal factors. A d. f. $G(x)$ is called *"infinitely divisible"* if for any $n=2, 3, \ldots$ there can be found a d. f. $F(x)$ such that $\overset{1}{F}(x)*\overset{2}{F}(x)*\ldots*\overset{n}{F}(x)=G(x)$. (It may be remarked, that $F(x)$ is not uniquely determined by this condition: in § 2 an example will be given that there exist two different distributions Φ_1 and Φ_2 with $\Phi_1*\Phi_1=\Phi_2*\Phi_2$). Theorems on infinitely divisible distributions of fundamental importance have been proved by A. KOLMOGOROFF[4]), A. KHINTCHINE[5]) and P. LÉVY[6]), but there are still many unsolved problems.[7]) Factorization problems are in general rather difficult, owing to the fact that in the algebra of distributions there is no cancellation law (i. e. it may occur that $F_1*F_2=F_1*F_3$ but $F_2\neq F_3$; this has been remarked first by KHINTCHIN[5]); an example will be given in § 2). In this connection it may be mentioned that no distribution, except the translations of the unity distribution (i. e. except the d. f. $E(x-m)$ where m is an arbitrary real number) has an inverse; by other words if $F(x)$ is a given d. f. there exists a d. f. $F^*(x)$ satisfying $F(x)*F^*(x)=E(x)$ if and only if $F(x)=E(x-m)$. This can be proved for instance as follows: let $a(t)$ and $b(t)$ denote measurable real functions defined in the interval $(0,\ 1)$ having $F(x)$ resp $F^*(x)$ as their d. f.-s; for example we may take for $a(t)$ resp. $b(t)$ the inverse functions (in the ordinary sense) of $F(x)$ resp. $F^*(x)$. Let us define in the unit square $(\leqq u\leqq 1; 0\leqq v\leqq 1)$ the functions $a(u,v)=a(u)$, $b(u,v)=b(v)$. The functions $a(u,v)$ resp. $b(u,v)$

---

[2]) H. CRAMÉR: Über eine Eigenschaft der normalen Verteilungsfunction. *Math. Zeitschrift* **41** (1936), 405–414.

[3]) D. RAIKOV: On the composition of Poisson laws, *Bull. Acad. Sci. URSS*, Ser. Math. (1938), 91–120.

[4]) A. KOLMOGOROFF: Sulla forma generale di un processo stocastico omogeneo, *Atti Accad. naz. Lincei, Rend.* **15** (1932), 805–808 and 866–869.

[5]) A. KHINTCHINE: Contribution à l'arithmétique des lois de distribution, *Bull. Math. Univ. Moscou* **1** (1937), 1–31.

[6]) P. LÉVY: Théorie de l'addition des variables aléatoires, (Paris, 1937) and Processus stochastiques et mouvement Brouwnien, (Paris, 1948).

[7]) Cf. H. CRAMÉR: Problems in probability theory, *Annals of Math. Statistics* **18** 1947), 165–193.

have (as functions of two variables in the unit square), the d. f.-s $F(x)$ resp. $F^*(x)$, further $a(u,v)$ and $b(u,v)$ are clearly independent. Let us apply now a measure preserving transformation of the unit square into the interval $[0, 1]$; $a(u,v)$ and $b(u,v)$ are thus transformed into two functions — say $A(t)$ and $B(t)$ — which have the d. f.-s $F(x)$ resp. $F^*(x)$ and which are evidently independent. It follows that the function $C(t)=A(t)+B(t)$ has the d. f. $F(x)*F^*(x)=E(x)$ i. e. $C(t)$ is equal to 0 almost everywhere, which contradicts the independence of $A(t)$ and $B(t)$ except if $A(t)$ is equal to a constant m almost everywhere, i. e. except if $F(x)=E(x-m)$ (in the latter case we have $F^*(x)=E(x+m)$) which proves our assertion. The idea of the above proof is due to F. RIESZ, to whom I am thankfull for some valuable remarks.

In the present paper we shall consider some factorization problems of the following type. Let $G(x)$ denote a d. f. and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ denote an n-tuple of real numbers which are different from 0. If there exists a d. f. $F(x)$ such that if $\xi_1, \xi_2, \ldots, \xi_n$ are independent randon variables all having the d. f. $F(x)$ then the randon variable η defined by

$$\eta = \lambda_1\xi_1 + \lambda_2\xi_2 + \ldots + \lambda_n\xi_n \tag{4}$$

has the d. f. $G(x)$, we shall say that $G(x)$ is $(\lambda_1, \lambda_2, \ldots, \lambda_n)$-divisible and shall write $G=F_{(\lambda_1,\lambda_2,\ldots\lambda_n)}$ or also $F=G/[\lambda_1, \lambda_2, \ldots \lambda_n]$. In this connection the following questions are of interest:

A) If $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is an arbitrary n-tuple of real numbers $(\lambda_k \neq 0; k=1,2\ldots., n)$ we may ask which distributions are $(\lambda_1, \lambda_2, \ldots, \lambda_n)$-divisible? The family of all $(\lambda_1, \lambda_2, \ldots, \lambda_n)$-divisible distributions shall be denoted by $F[\lambda_1, \lambda_2, \ldots, \lambda_n]$.

B) If the d. f. $G(x)$ belongs to the family $F[\lambda_1, \lambda_2, \ldots, \lambda_n]$ we may ask whether $F(x)=G(x)/(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is uniquely determined or not?

C) If $G(x)$ belongs to the family $F(\lambda_1, \lambda_2, \ldots, \lambda_n)$ further if we know already that $F=G/(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is uniquely determined, we may ask how to calculate $F(x)$ or some of its parameters (its mean value, its standard deviation, its moments or semi-invariants, etc) explicitly?

It may seem at the first sight that when considering the above questions one has to begin by solving question **A)** and then proceed to questions **B)** and **C)**, which seem to be of secondary importance compared with question **A)**. As a matter of fact this is true from a purely logical point of view; on the other hand from the point of view of the mathematical statistician questions **B)** and **C)** are of primary importance while question **A)** does not occur generally at all. As a matter of fact, from the point of view of statistics our problem may be described as that of determining an unknown distribution by means of *indirect observations*, i. e. observations bearing not on a single variable distributed according to the unknown distribution in question, but on the sum $\eta=\lambda_1\xi_1+\lambda_2\xi_2+\ldots+\lambda_n\xi_n$ where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are known real numbers while $\xi_1, \xi_2, \ldots, \xi_n$ are independent variables having all the same

(unknown) distribution. In this case the existence of the unknown distribution is beyond doubt, while the questions of unicity and explicit calculation present themselves as important probleme: we have to know whether our observations concerning η are sufficient to determine the distribution in question uniquely, because if not, we have to change the design of our experiment (i. e. the values of the λ_k-s). If an appropriate set of λ_k-s has been chosen (i. e. if the unicity is ensured) we have to calculate effectively the parameters of the unknown distribution.

In what follows we shall consider in the first place questions **B)** and **C)**. In § 1 we shall restrict ourselves to the case when the variable η defined by (4) is bounded. In this case — which is the most important from the practical point of view — problems **B)** and **C)** are relatively simple and can be solved completely. In § 2 we consider the general case (η unbounded). In § 3 the problem will be generalized; we consider the simultaneous determination of more unknown distributions.

In spite of what has been said above, question **A)** is also of considerable theoretical interest, but seems to be very difficult. We add only some simple remarks. First of all it is clear that any family $F(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is a *subalgebra* of the algebra of all distributions: that is to say if G_1 and G_2 both belong to $F(\lambda_1, \lambda_2, \ldots, \lambda_n)$ so does $G_1 * G_2$. As a matter of fact, we have

$$(5) \qquad (G_1 * G_2)/(\lambda_1, \lambda_2, \ldots, \lambda_n) = G_1/(\lambda_1, \ldots, \lambda_n) * G_2/(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

By other words, $F(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is also a semi-group. It is interesting to compare these subalgebras to two well-known subalgebras of the algebra of distributions: the family of *stable* distributions F_s and the family of *infinitely divisible* distributions F_i. A d. f. $F(x)$ is called stable, if for any $a > 0, \alpha > 0$ and real b and β the d. f. $F(ax+b) * F(\alpha x+\beta)$ is also of the form $F(Ax+B)$ with $A > 0$. By other words, a d. f. $F(x)$ is called stable if the family of all distributions of the form $F(ax+b)$ $(a > 0)$ form a subalgebra (with respect to the composition defined by (1)). Let further F_0 denote the family of distributions which are contained in every $F(\lambda_1, \lambda_2, \ldots, \lambda_n)$ where $\sum_{k=1}^{n} \lambda_k \neq 0$; F_0 is also a subalgebra. Now it follows simply that F_s is contained in F_0, whereas F_0 is contained in F_i. The second assertion is trivial, while the first follows by explicit calculation using the results of KHINTCHINE and LÉVY[8]) on the general form of stable distributions. That F_i is not identical with F_0 follows from the remark that the Poisson distribution, which belongs to F_i belongs to $F(\lambda_1, \lambda_2, \ldots, \lambda_n)$ if and only if all λ_k-s are equal. The significance of the restriction $\sum_{k=1}^{n} \lambda_k \neq 0$ used in the definition of F_0 will be clear form the results of § 1.

[8]) A. KHINTCHINE and P. LÉVY: Sur les lois stables, *Comptes Rendus Acad. Sci.* (Paris) **202** (1936), 374—376.

§ 1. Bounded variables.

Theorem 1. *Let us suppose that $G(x)$ ist the d. f. of a bounded variable (i. e. $G(x)$ is constant outside a finite interval) further that $G(x)$ belongs to $F(\lambda_1, \lambda_2, \ldots, \lambda_n)$. It follows, that $F(x) = G(x)/(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is uniquely determined if and only if $\sum_{k=1}^{n} \lambda_k \neq 0$.*

Proof. Let us put $\Lambda_s = \sum_{k=1}^{n} \lambda_k^s$; $s = 1, 2, 3, \ldots$. The necessity of the condition $\sum_{k=1}^{n} \lambda_k \neq 0$ is clear, because if $\Lambda_1 = 0$ and $\eta = \sum_{k=1}^{n} \lambda_k \xi_k$ it follows $\eta = \sum_{k=1}^{n} \lambda_k (\xi_k + a)$ where a is any real constant; thus if $F(x) = G(x)/(\lambda_1, \lambda_2, \ldots, \lambda_n)$ it follows that also $F(x-a) = G(x)/(\lambda_1, \lambda_2, \ldots, \lambda_n)$ for any real a, and thus $G(x)/(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is not uniquely determined. To prove the sufficiency of the condition $\Lambda_1 \neq 0$ we introduce the characteristic functions (abbreviated: c. f.) of $F(x)$ and $G(x)$ by putting

$$(6) \qquad f(t) = \int_{-\infty}^{+\infty} e^{ixt}\, dF(x) \qquad g(t) = \int_{-\infty}^{+\infty} e^{ixt}\, dG(x) \qquad (t \text{ real}).$$

By a well known property of the Fourier-transform it follows that the relation $F(x) = G(x)/(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is equivalent to

$$(7) \qquad g(t) = f(\lambda_1 t) f(\lambda_2 t) \ldots f(\lambda_n t).$$

As there is a one-one correspondence between the d. f. and its c. f. the sufficiency of the condition $\Lambda_1 \neq 0$ in *Theorem 1.* is equivalent to the following statetement: If $\Lambda_1 \neq 0$ and $g(t)$ is the c. f. of a bounded variable, there exists not mere than one c. f. $f(t)$ which solves the functional equation (7). It is easy to see that the c. f. of a bounded variable is an entire function of exponential type.[9]) As further it is evident, that the boundedness of $\eta = \sum_{k=1}^{n} \lambda_k \xi_k$ implies the boundedness of the variables ξ_k, it follows that in (7) both $f(t)$ and $g(t)$ are analytic functions of exponential type. As $f(0) = g(0) = 1$ there exists an $r > 0$ such that $f(t) \neq 0$ and $g(t) \neq 0$ in the circle $|t| \leqq r$ of the complex plane. Introducing the functions $\varphi(t) = \log f(t)$ and $\gamma(t) = \log g(t)$ it follows that both $\varphi(t)$ and $\lambda(t)$ are analytic in $|t| \leqq r$ (we choose those branches of $\log f(t)$ resp. $\log g(t)$ which take the value 0 for $t = 0$). Now we obtain from (7)

---

[9]) Cf. G. Pólya: Remarks on characteristic functions. *Proceedings of the Berkeley Symposium on Mathematical Staatistics and Probability,* Berkeley and Los Angeles (1945—46) 115—123.

(8) $$\gamma(t)=\sum_{k=1}^{n}\varphi(\lambda_k t).$$

Differentiating (8) s times ($s=1,2,\ldots$) and substituting $t=0$ we obtain

(9) $$\gamma^{(s)}(0)=A_s\varphi^{(s)}(0).$$

It follows that $\varphi^{(s)}(0)$ is determined for those values of s for which $A_s\neq 0$. Now we use the following elementary.

Lemma 1. *Let $\lambda_1,\lambda_2,\ldots,\lambda_n$ denote arbitrary real numbers and put $A_s=\sum_{k=1}^{n}\lambda_k^s$ for $s=1,2,\ldots$. If $A_1\neq 0$, the set of those values of s for which $A_s=0$ is finite.*

To prove Lemma 1 we remark that for s even evidently $A_s\neq 0$, and for s odd we may put $A_s=\sum_{j=1}^{k'}A_j\mu_j^s$ where $\mu_1>\mu_2>\ldots>\mu_{k'}>0$ and the A_j-s are integers $\neq 0$. It follows $\frac{A_s}{\mu_1^s}\to A_1\neq 0$ for $s\to\infty$ which proves our Lemma. Applying Lemma 1, if $f_1(t)$ and $f_2(t)$ are two solutions of the functional equation (7), putting $\varphi_1(t)=\log f_1(t)$ and $\varphi_2(t)=\log f_2(t)$ it follows that $\varphi_1^{(s)}(0)=\varphi_2^{(s)}(0)$ for $s=1,2,\ldots$ except for a finite set of exceptional values of s; thus $\varphi_1(t)-\varphi_2(t)=P(t)$ is a polynomial in t, and thus $f_1(t)=f_2(t)e^{P(t)}$. As both $f_1(t)$ and $f_2(t)$ are entire functions of exponential type, the polynomial $P(t)$ must be of order $\leqq 1$. But as $A_1\neq 0$, $s=1$ is not exceptional and $P(t)$ must be a constant; finally as $f_1(0)=f_2(0)=1$ it follows $P(t)\equiv 0$ and therefore $f_1(t)\equiv f_2(t)$ which proves the sufficiency of the condition $A_1\neq 0$, and thus proves *Theorem 1*. The above proof furnishes also a complete solution of question C) in the special case in which it is supposed that $A_s\neq 0$ for $s=1,2,\ldots$. We obtain

Theorem 2. *Let $G(x)$ denote the d. f. of a bounded variable and let us suppose that $G(x)$ belongs to $F(\lambda_1,\lambda_2,\ldots,\lambda_n)$ where it is supposed that $A_s\neq 0$ for $s=1,2,3,\ldots$. Let $g(t)$ denote the c. f. of $G(x)$ and $f(t)$ the c. f. of $F(x)=G(x)/(\lambda_1,\lambda_2,\ldots,\lambda_n)$ (the latter being uniquely determined according of Theorem 1) further let us put*

(10) $$\log g(t)=\sum_{s=1}^{\infty}\frac{g_s(it)^s}{s!}.$$

It follows that

(11) $$\log f(t)=\sum_{s=1}^{\infty}\frac{g_s(it)^s}{A_s\,.\,s!}.$$

Thus if the semi-invariants g_s of $G(x)$ are known, the semi-invariants f_s of $F(x)$ can be simply obtained by the formula $f_s=\frac{g_s}{A_s}$. From the point of view of the statistician it is important to calculate the moments

$$(12)\qquad \alpha_s = \int_{-\infty}^{+\infty} x^s dF(x) \qquad s = 1, 2, \ldots$$

of the d. f. $F(x)$. Putting

$$(13)\qquad \beta_s = \int_{-\infty}^{+\infty} x^s dG(x) \qquad s = 1, 2, \ldots$$

we obtain by using well known relations between the moments and the semi invariants the expression of the α_s by means of the β_s:

$$(14)\begin{cases} \alpha_1 = \dfrac{\beta_1}{A_1} \\ \alpha_2 = \dfrac{\beta_2}{A_2} + \beta_1^2\left(\dfrac{1}{A_1^2} - \dfrac{1}{A_2}\right) \\ \alpha_3 = \dfrac{\beta_3}{A_3} + 3\beta_2\beta_2\left(\dfrac{1}{A_1A_2} - \dfrac{1}{A_3}\right) + \beta_1^3\left(\dfrac{1}{A_1^3} - \dfrac{3}{A_1A_2} + \dfrac{2}{A_3}\right) \\ \alpha_4 = \dfrac{\beta_4}{A_4} + 4\beta_1\beta_2\left(\dfrac{1}{A_1A_3} - \dfrac{1}{A_4}\right) + 3\beta_2^2\left(\dfrac{1}{A_2^2} - \dfrac{1}{A_4}\right) + 6\beta_1^2\beta_2\left(\dfrac{1}{A_1^2A_2} - \dfrac{1}{A_2^2} - \dfrac{2}{A_1A_3} + \dfrac{2}{A_4}\right) + \\ \vdots \\ \text{etc.} \qquad + \beta_1^4\left(\dfrac{1}{A_1^4} - \dfrac{6}{A_1^2A_2} + \dfrac{3}{A_2^2} + \dfrac{3}{A_1A_3} - \dfrac{6}{A_4}\right) \end{cases}$$

Theorem 2 includes the most important case, in which all $\lambda_k > 0$. On the other hand if only $A_1 \neq 0$ is supposed, we can calculate as above the semi-invariants f_s of $F(x)$ for those values s for which $A_s \neq 0$, but for the exceptional values of s the above proof gives no method for calculating f_s.

We add still some remarks on the case $A_1 = 0$. We have seen that in this case if there exists a solution $F(x)$ there is an infinity of solutions, viz. $F(x-a)$ fon any real a One may think that these are all solutions, i. e. that $F(x)$ is determined up to a translation, but this is not true. For example let us consider the case $\lambda_1 = 1$, $\lambda_2 = -1$, that is we put $\eta = \xi_1 - \xi_2$ and let us suppose that η takes the values: $-2, -1, 0, +1, +2$ the corresponding probabilites being: $\frac{1}{45}, \frac{2}{9}, \frac{23}{45}, \frac{2}{9}, \frac{1}{45}$; the following solutions for the independent variables ξ_1, ξ_2 are possible: ξ_1, ξ_2 take the values a, $a+1$, $a+2$ with the corresponding probabilities: $\frac{5 \mp \sqrt{5}}{30}, \frac{2}{3}, \frac{5 \pm \sqrt{5}}{30}$ or withthe corresponding probabilities: $\frac{5 \mp 2\sqrt{5}}{15}, \frac{1}{3}, \frac{5 \pm 2\sqrt{5}}{15}$. Thus up to translations we have four distinct solutions for $F(x)$. But among the different solutions of our problem (in the case $A_1 = 0$) there can be at most one d. f. which is symmetric, i. e. which satisfies the condition $F(x) + F(-x) = 1$. As a matter of fact, if the independent random variables ξ_k $k = 1, 2, \ldots, n$ have a symmetric d. f. $F(x)$, the sum $\eta_1 = \sum_{k=1}^{n} \lambda_k \xi_k$ has the same d. f. as the sum $\eta_2 = \sum_{k=1}^{n} |\lambda_k| \xi_k$, because

ξ_k and $-\xi_k$ have the same d. f., and thus we may suppose that every $\lambda_k > 0$, in which case Theorem 1 ensures the unicity of $F(x)$. Evidently a symmetric solution $F(x)$ can exist only if $G(x)$ is symmetric. But even if there is a symmetric solution, there may exist other asymetric solutions also, as it is shown by the following example: $\eta = \xi_1 - \xi_2$ takes the values $-2, -1, 0, +1, +2$ with the corresponding probabilities $\frac{4}{81}, \frac{20}{81}, \frac{33}{81}, \frac{20}{81}, \frac{4}{81}$ there exists a symmetric solution: ξ_1, ξ_2 take the values $-1, 0, +1$ with the corresponding probabilites: $\frac{2}{9}, \frac{5}{9}, \frac{2}{9}$; but besides this there exists an asymetric sokutiin also: ξ_1, ξ_2 take the values $-1, 0, +1$ (or the values a, $a+1$, $a+2$ with arbitrary real a) with the corresponding probabilities: $\frac{4}{9}, \frac{4}{9}, \frac{1}{9}$. These examples are characteristic for the situation when $A_1 = 0$.

§ 2. The general case.

We begin by remarking that in case it is not supposed that $G(x)$ is the d. f. of a bounded variable, but only that the c. f. $g(t)$ of $G(x)$ is analytic in some interval $-r \leqq t \leqq +r$ $(r > 0)$ and if $A_s \neq 0$ for $s = 1, 2, \ldots$ the assertions of Theorems 1. and 2. remain valid.

Theorem 3. *If the d. f. $G(x)$ belongs to $F(\lambda_1, \lambda_2, \ldots, \lambda_n)$ where $A_s = \sum_{k=1}^{n} \lambda_k^s \neq 0$ for $s = 1, 2, \ldots$, further if the c. f. $g(t)$ of $G(x)$ is analytic in an interval $-r \leqq t \leqq +r$ $(r > 0)$, it follows that $F(x) = G(x) / \lambda_1, \lambda_2, \ldots, \lambda_n)$ is uniquely determined and its characteristic function $f(t)$ can be calculated as described in Theorem 2.*

To prove Theorem 3. it suffices to remark that if $g(t)$ is analytic in $(-r, +r)$, it is analytic on the whole real axis accordig to a theorem of D. A. RAIKOV[10]; thus it follows, by a theorem of S. MAZURKIEWICZ[11] that $f(t)$ is also analytic in the same domain, and therefore we obtain, exactly as in § 1 that denoting by f_s resp. g_s the semi-invariants of $F(x)$ resp. $G(x)$ we have $f_s = \frac{g_s}{A_s}$. It follows that $f(t)$ is uniquely determined in a neighborhood of the origin and hence, by analytic continuation, on the whole real axis. This theorem includes a special case of the theorem of H. CRAMÉR[2]: if $G(x)$ is a normal distribution so is $F(x)$: as a matter of fact the c. f. of $G(x)$ is

[10] Cf: A theorem from the theory of analytic characteristic functions, *Bull. Inst. Math. Méc. Univ. Towsk*, **2** (1938), 8–11; the theorem has been recently rediscovered by R. P. BOAS: Sur les séries et intégrales de Fourier a coeficients positifs, *Comptes Rendus Acad. Sci.* (Paris) **228** (1949), 1837–38.

[11] S. MAZURKIEWICZ: Un théoreme sur les fonctions caractéristiques, *Bull. Int. Acad. Polon. Cl. Sci. Math.* 1940–1946, (1948), 1–3.

in this case $g(t)=im\,t-\frac{\sigma^2 t^2}{2}$ this means $g_s=0$ for $s=3,4,\ldots$ and therefore $f_s=0$ for $s=3,4,\ldots$.

Now let us consider the case when nothing is supposed about the analytic character of $g(t)$ It is easy to find examples showing that in this case $F(x)$ is not always unique. Let us consider the most simple case: $n=2, \lambda_1=\lambda_2=1$. In this case it suffices to show that there exist two different c. f.-s $f_1(t)$ and $f_2(t)$ for which $f_1^2(t)=f_2^2(t)$ is identically valid. Such a pair may be constructed as follows: let $f_1(t)$ denote the c. f. of a randon variable taking the values $\pm(2k+1)$ $(k=0,1,2,\ldots)$ with the corresponding probabilities $\frac{4}{\pi^2(2k+1)^2}$.

It follows

$$f_1(t)=\frac{8}{\pi^2}\sum_{k=0}^{\infty}\frac{\cos(2k+1)t}{(2k+1)^2}=1-\frac{2|t|}{\pi}$$

for $-\pi\leqq t\leqq\pi$ and is periodic with period 2π. Let further $f_2(t)$ denote the c. f. of a random variable taking the value 0 with probability $\frac{1}{2}$ and the values $\pm(4k+2)$ $(k=0,1,2,\ldots)$ with the corresponding probabilities $\frac{2}{\pi^2(2k+1)^2}$ and thus

$$f_2(t)=\frac{1}{2}+\frac{4}{\pi^2}\sum_{k=0}^{\infty}\frac{\cos(4k+2)t}{(2k+1)^2}=1-\frac{2|t|}{\pi}$$

for $-\frac{\pi}{2}\leqq t\leqq\frac{\pi}{2}$ and is periodic with period π. Clearly $f_2(t)=|f_1(t)|$ and thus $f_2^2(t)=f_1^2(t)$, identically.

Using the above example it is easy to show that there is no cancellation law in the semi-group of distributions. Let $F_1(x)$ and $F_2(x)$ denote the d. f.-s corresponding to $f_1(t)$ and $f_2(t)$. Evidently $f_3(t)=\frac{1}{2}(f_1(t)+f_2(t))$ is also a c. f, the corresponding d. f. shall be denoted by $F_3(x)$. As we have $f_3(t)f_1(t)=f_3(t)f_2(t)$ identically, it follows $F_3(x)*F_1(x)=F_3(x)*F_2(x)$ but $F_1(x)\not\equiv F_2(x)$.

Analysing the above example we see that, putting $f_1^2(t)=f_2^2(t)=g(t)$, between two consecutive roots of $g(t)f_1(t)$ and $f_2(t)$ are equal alternatively to the same resp. to different branches of $\sqrt{g(t)}$. This phenomenon can occur only if $g(t)$ has real roots; if $g(t)\neq 0$ for real t and $f^2(t)=g(t)$ (or more generally if $f^n(t)=g(t)$ with $n\geqq 2$) $f(t)$ is uniquely determined as that branch of $\sqrt[n]{g(t)}$ which is equal to 1 for $t=0$. We proceed to prove that the condition $g(t)\neq 0$ for real t ensures the unicity under fairly general conditions.

Theorem 4. *Let us suppose that the n-tuple of real numbers* $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ *satisfies the following conditions:*

a) $\lambda_1 = \lambda_2 = \ldots = \lambda_{n_1} \neq 0$ $(n_1 \geqq 1)$ *and* $|\lambda_{n_1+r}| < |\lambda_1|$ *for* $r = 1, 2, \ldots, n_2$; $n_2 = n - n_1$.

b) *Putting* $\mu_r = \frac{\lambda_{n_1+r}}{\lambda_1}$; $(r = 1, 2, \ldots n_2)$

$$\frac{1}{n_1} \sum_{r=1}^{M_2} |\mu_r|^p < 1$$ *holds, where p is a positive integer.*

c) *Putting* $A_s = \sum_{k=1}^{n} \lambda_k^s$, $A_s \neq 0$ *for* $s = 1, 2, \ldots (p-1)$.

Let us suppose further that the d. f. $G(x)$ *satisfies the following conditions:*

d) $G(x)$ *belongs to* $F(\lambda_1, \lambda_2, \ldots, \lambda_n)$

e) *the absolute moments* $B_s = \int_{-\infty}^{+\infty} |x|^s dG(x)$ *of order* $s = 1, 2, \ldots, p$ *of* $G(x)$ *exist.*

f) *denoting by* $g(t)$ *the c. f. of* $G(x)$ *we have* $g(t) \neq 0$ *for real t.*

It foliows that $F(x) = G(x)/(\lambda_1, \lambda_2, \ldots, \lambda_n)$ *is uniquely determined.*

Before proving this theorem, let us mention some of its consequences:

Corollary 1. *If* $n = 2$, $\lambda_1, + \lambda_2 \neq 0$ *further* $\int_{-\infty}^{+\infty} |x| dG(x)$ *exists and* $g(t) = \int_{-\infty}^{+\infty} e^{itx} dG(x) \neq 0$ *for real t, then if* $G(x)$ *belongs to* $F(\lambda_1, \lambda_2)$, $F(x) = G(x)/(\lambda_1, \lambda_2)$ *is uniquely determined.*

We shall prove later (Theorem 5) that the existence of $\int_{-\infty}^{+\infty} |x| dG(x)$ is not necessary for the validity of this Corollary. Corollary 1 follows from Theorem 4, because either $\lambda_1 = \lambda_2$ in which case $n_2 = 0$, $n_1 = 2$ and conditions a, b, c, are trivially satisfied, or $\lambda_1 \neq \lambda_2$ and thus, with respect to $\lambda_1 + \lambda_2 \neq 0$, $|\lambda_1| \neq |\lambda_2|$, and therefore either $|\lambda_1| > |\lambda_2|$ or $|\lambda_1| < |\lambda_2|$; in the first case conditions a, b, and c are satisfied; the latter case can be reduced to the first by interchanging λ_1 and λ_2.

Corollary 2. *If the moments of any order of* $G(x)$ *exist, condition* b) *may be omitted.*

As a matter of fact, if condition a) is satisfied, it follows that condition b) is also satisfied if p is a sufficiently large integer, and thus Theorem 4 can be applied.

Finally it is clear that if $p = 1$, condition c) contains no restriction on the A_k-s, but in this case condition b) implies that $A_1 \neq 0$. As a matter of fact, we have

$$|\Lambda_1| = n_1 |\lambda_1| \left| 1 + \frac{1}{n_1} \sum_{r=1}^{n_2} \mu_r \right| \geqq n_1 |\lambda_1| \left(1 - \frac{1}{n_1} \sum_{r=1}^{n_2} |\mu_r| \right) > 0.$$

Proof Theorem 4. Let us suppose that there exist two different solutions $F_1(x)$ and $F_2(x)$ and let us denote their c. f.-s by $f_1(t)$ resp. $f_2(t)$. Let us put $\varphi(t) = \log \frac{f_1(t)}{f_2(t)}$. According to f), $g(t) \neq 0$ which implies that neither $f_1(t)$ nor $f_2(t)$ vanish for t real, and thus $\varphi(t)$ is a continuous function. According to our hypotheses we have

$$\sum_{k=1}^{n} \varphi(\lambda_k t) = 0. \tag{15}$$

Now according to e) the absolute moments of order s, $(s = 1, 2, \ldots p)$, of $G(x)$ exist, and it follows[12]) that the same holds for both $F_1(x)$ and $F_2(x)$. According to a well known theorem[13]) we obtain that $f_1(t)$ and $f_2(t)$ are p times derivable at $t = 0$: as a matter of fact we have

$$f_j^{(s)}(0) = i^s \int_{-\infty}^{+\infty} x^s dF_j(x) \qquad (j = 1, 2). \tag{16}$$

As in the proof of Theorem 1. we conclude making use of c), that the first p derivatives of $\log f_1(t)$ and $\log f_2(t)$ take the same values for $t = 0$. It results that the first p derivatives of $\varphi(t)$ vanish at $t = 0$ and therefore (with respect to $\varphi(0) = 0$)

$$\varphi(t) = o(t^p). \tag{17}$$

Now substituting $t = \frac{u}{\lambda_1}$ we obtain from (15)

$$\varphi(u) = -\frac{1}{n_1} \sum_{r=1}^{n_2} \varphi(\mu_r u). \tag{18}$$

Applying (18) N times $(N = 1, 2, 3, \ldots)$ we have

$$\varphi(u) = \frac{(-1)^N}{n_1^N} \sum_{\substack{h_1 + h_2 + \ldots + h_{n_2} = N \\ h_r \geqq 0;\, r = 1, 2, \ldots n_2}} \varphi(\mu_1^{h_1} \mu_2^{h_2} \ldots \mu_{n_2}^{h_{n_2}} u). \tag{19}$$

Using (17) it follows

$$|\varphi(u)| = 0\left(u^p \left(\frac{1}{n_1} \sum_{r=1}^{n_2} |\mu_r|^p \right)^N \right). \tag{20}$$

Making use of condition b) this gives $\varphi(u) = 0$ for any real u, which proves our theorem.

The importance of condition b) is clear from the above proof. Note that in case $n_1 \geqq n_2$ condition b) is automatically satisfied.

Now let us consider the case $n = 2$. We prove

---

[12]) Mazurkiewicz, l. c.[11])

[13]) Cf. H. Cramér: Mathematical methods of statistics, Princeton, (1946), p. 89.

Theorem 5. *Let us suppose* $n=2$, $\lambda_1+\lambda_2\neq 0$. *If* $G(x)$ *belongs to* $F(\lambda_1,\lambda_2)$ *and the c. f.* $g(t)$ *of* $G(x)$ *does not vanish on the real axis,* $F(x)=G(x)/(\lambda_1,\lambda_2)$ *is uniquely determined.*

Proof: The case $\lambda_1=\lambda_2$ is trivial, as it has been remarked on p. 144. If $\lambda_1\neq\lambda_2$, we may suppose $|\lambda_1|<|\lambda_2|$. Let us put $\lambda=\frac{\lambda_1}{\lambda_2}$. If we define $\varphi(t)$ as in the proof of Theorem 4 it follows

$$\varphi(t)=-\varphi(\lambda t).$$

Thus

$$\varphi(t)=(-1)^n\varphi(\lambda^n t)$$

and by virtue of $|\lambda|<1$ it follows $\varphi(t)\equiv 0$.

The conditions of Theorem 4 exclude the case in which, putting $\lambda=\max\limits_{1\leqq k\leqq n}|\lambda_k|$ we have $\lambda_{r_1}=\lambda$ and $\lambda_{r_2}=-\lambda$ for some r_1 and r_2. We now prove a theorem including this case also. To simplify the notations let us introduce the following definition: an n-tuple of real numbers $(\lambda_1,\lambda_2,\ldots\lambda_n)$ is called *p-regular* if $\lambda_1=\lambda_2=\ldots=\lambda_{n_1}=\lambda\neq 0$, $\lambda_{n_1+1}=\lambda_{n_1-2}=\ldots=\lambda_{n_1+n_2}=-\lambda$; $n_1\neq n_2$; further putting $\mu_r=\frac{\lambda_{n_1+n_2+r}}{\lambda}$ $r\equiv 1,2,\ldots n_3$, $(n_3=n-n_1-n_2)$, we have

$$\frac{1}{|n_1-n_2|}\sum_{r=1}^{n_3}|\mu_r|^p<1 \tag{21}$$

where p is a positive integer. We prove now

Theorem 6. *Let us suppose that* $(\lambda_1,\lambda_2,\ldots\lambda_n)$ *is a p-regular n-tuple of real numbers, further that conditions* c), d), e), *and* f) *of Theorem 4. are satisfied. It follows that* $F(x)=G(x)/(\lambda_1,\lambda_2,\ldots\lambda_n)$ *is uniquely determined.*

Proof. Let us define $\varphi(t)$ as in the proof of Theorem 4. We obtain

$$n_1\varphi(u)+n_2\varphi(-u)+\sum_{r=1}^{n_3}\varphi(\mu_r u)=0. \tag{22}$$

Putting $\varphi(u)+\varphi(-u)=\psi(u)$, applying (22) with $-u$ instead of u and adding the indentity thus obtained to (22) it follows

$$\psi(u)=-\frac{1}{(n_1+n_2)}\sum_{r=1}^{n_3}\psi(\mu_r u). \tag{23}$$

As we supposed (21), $\frac{1}{n_1+n_2}\sum\limits_{r=1}^{n_3}|\mu_r|^p<1$ holds a fortiori.

Using the same argument as in the proof of Theorem 4 we obtain $\psi(u)\equiv 0$, and thus $\varphi(-u)=-\varphi(u)$. Substituting $-\varphi(u)$ in place of $\varphi(-u)$ in (22) it follows

$$(n_1-n_2)\varphi(u)+\sum_{r=1}^{n_3}\varphi(\mu_r u)=0. \tag{24}$$

Now we apply the above method again, and obtain by virtue of (21) that $\varphi(u)\equiv 0$.

If $n_1 = n_2$ (n_1 and n_2 having the same meaning as in the definition of p-regular n-tuples), we can not apply Theorem 6. Nevertheless in this case we can also conclude that $G(x)/(\lambda_1, \ldots \lambda_n)$ is uniquely determined, provided that $(\mu_1, \mu_2, \ldots \mu_{n3})$ is p-regular. As a matter of fact, if $n_1 = n_2$, (24) gives

$$\sum_{r=1}^{n_3} \varphi(\mu_r u) = 0 \tag{25}$$

and thus if $(\mu_1, \mu_2, \ldots \mu_r)$ is p-regular, Theorem 6 can be applied. For example if $n_3 = 1$, unicity is always ensured.

§ 3. More unknown distributions.

In this § we consider the following problem: let ξ_k $(k = 1, 2, \ldots, n)$ denote independent random variables, the unknown distributions of which are not supposed to be equal. To determine the n unknown distributions, the knowledge of the distribution of n different linear forms

$$\eta_j = \sum_{k=1}^{n} c_{jk}\xi_k \qquad (j = 1, 2, \ldots n) \tag{26}$$

is necessary, and under appropriate conditions, also sufficient. We introduce some definitions, which are analogous to those employed in the preceding sections. Let us denote the d. f. of ξ_k by $F_k(x)$ and the d. f. of η_j defined by (26) by $G_j(x)$ $(k, j = 1, 2, \ldots n)$. Let us denote the matrix (c_{ik}) by Λ. If (26) holds, we shall say that the system $(G_1, G_2, \ldots G_n)$ is Λ-divisible, and shall write $(F_1, F_2, \ldots F_n) = (G_1, G_2, \ldots G_n)/\Lambda$. The class of all Λ-divisible systems of n d. f.-s shall be denoted by $F(\Lambda)$.

Theorem 7. *If the system of distributions $(G_1, G_2, \ldots G_n)$ belongs to $F(\Lambda)$ where $\Lambda = (c_{jk})$, further if the c. f. $g_1(t)$ of $G_1(x)$ is an entire function of order $\leqq 2$, it follows that the system $(F_1, F_2, \ldots F_n) = (G_1, G_2, \ldots G_n)/\Lambda$ is uniquely determined if and only if the determinants $D_1 = |c_{jk}|$ and $D_2 = |c_{jk}^2|$ are different from zero.*

Proof. First of all, if $g(t)$ is an entire function of order $\leqq 2$, the same holds for the c. f. $f_k(t)$ of $F_k(x)$ for $k = 1, 2, \ldots n$, (according to the theorem of MAZURKIEWICZ[11])), and thus also for the c. f. $g_j(t)$ of $G_j(x)$ for $j = 2, 3, \ldots n$ also. To prove the sufficiency of the conditions $D_1 \neq 0$, $D_2 \neq 0$, we need the following lemmas:

Lemma 2. *Let $a_1, a_2, \ldots a_n$ and $\lambda_1, \lambda_2, \ldots \lambda_n$ denote real numbers, and let us put*

$$\Lambda_s = \sum_{k=1}^{n} a_k \lambda_k^s.$$

If $\Lambda_1 \neq 0$ and $\Lambda_2 \neq 0$, the set of those positive integer values of s, for which $\Lambda_s = 0$, is finite.

The proof of Lemma 2. is analogous to that of Lemma 1. and may be left to the reader.

Lemma 3. *Let $A=(c_{jk})$ denote an $n.n$ matrix of real numbers and let us consider the determinants*

$$D_s=|c_{jk}^s|.$$

If $D_1\neq 0$ and $D_2\neq 0$, it follows that the set of those positive integer values of s, for which $D_s=0$, is finite.[14])

The proof of Theorem 7 follows step by step the pattern of proof employed in proving Theorem 1, by using Lemma 3 instead of Lemma 1. In case D_s does not vanish for $s=1, 2, 3, \ldots$, the semi-invariants of $F_k(x)$ can be explicitly calculated in a similar way as the semi-invariants of $F(x)$ have been determined in Theorem 2.

The necessity of the condition $D_1\neq 0$ is readily seen. As a matter of fact, if $D_1=0$ and $(t_1, t_2, \ldots, t_n)$ is a non-trivial solution of the homogeneous system of equations $\sum_{k=1}^{n} c_{jk}t_k=0$ $(j=1, 2, \ldots, n)$ it follows that if the sequence of variables ξ_k satisfie (26) the sequence $\xi_k^*=\xi_k+t_k$ $(k=1, 2, \ldots, n)$ represents an other solution. To prove the necessity of the condition $D_2\neq 0$, let us suppose that for a given matrix $A=(c_{jk})$ we have $D_2=0$ Let σ_k denote a sequence of positive numbers, $\sigma_k\geqq h>0$, $(k=1, 2, \ldots, n)$. Let ξ_k denote a normally distributed random variable with parameters $(0, \sigma_k)$. It follows that η_j, defined by (26) is also normally distributed with parameters $(0, \varrho_j)$ where $\varrho_j=\sum_{k=1}^{n} c_{jk}^2\sigma_k^2$. If $(\tau_1, \tau_2, \ldots, \tau_n)$ denotes a non-trivial solution of the homogeneous system of equations $\sum_{k=1}^{n} c_{jk}^2\tau_k=0$, let us put $\sigma_k^*=\sqrt{\sigma_k^2+\frac{\tau_k h^2}{2\tau}}$ where $\tau=\max_{1\leq k\leq n}|\tau_k|$ (note that according to our suppositions $\sigma_k^2+\frac{\tau_k h^2}{2\tau}\geqq \geqq\sigma_k^2-\frac{h^2}{2}\geqq\frac{h^2}{2}>0$ for $k=1, 2, \ldots, n$). It follows that if ξ_k^* denotes a normally distributed random variable with parameters $(0, \sigma_k^*)$ and if we denote $\eta_j^*=\Sigma c_{jk}\xi_k^*$, $(j=1, 2, \ldots, n)$, the variables η_j^* are also normally distributed with parameters $(0, \varrho_j)$ $(j=1, 2, \ldots, n)$ which proves the necessity of the condition $D_2\neq 0$.

§ 4. Applications.

The necessity of the condition $D_2\neq 0$ in Theorem 7 is of a paradox character. As a matter of fact, if the variables η_j in (26) are given, the variables ξ_k are completely determined thereby, provided that $D_1\neq 0$; why do we need also $D_2\neq 0$? The answer is clear: if we choose two different

[14]) This lemma has been kindly suggested to me by Dr. I. Vincze.

sets of random variables (η_j) and (η_j^*), η_j and η_j^* having both the same d. f $G_j(x)$ $(j=1,2,\ldots,n)$ we obtain in both case a definite set of variables (ξ_k) resp. (ξ_k^*) but the d. f. of ξ_k and that of ξ_k^* may be different if $D_2 \neq 0$ is not supposed.

This rather curious phenomenon has interesting practical consequences. Let us consider for instance the weighing problem of HOTELLING.[15]) HOTELLING has shown that under suitable conditions, the weights of n individual objects may be determined more accurately by weighing the objects in combination (by a chemical balance for instance putting some of the objects in the first pan, and the others in the second pan) rather than weighing each one separately. To such a weighing design there corresponds a matrix (c_{jk}) where $c_{jk}=\pm 1$. Especially HOTELLING found that the most usefull experimental design corresponds to HADAMARD matrices in case for the given number n of objects suh a matrix exists.[16]) Theorem 7 shows, that if we are interested in determining the distribution according to their weights (or lengths, resistances, etc.) of *n classes of objects* (instead of n individual objects) caution is necessary: in fact only such matrices (c_{jk}) can be applied for wich $D_2=|c_{jk}^2| \neq 0$. This condition excludes for instance all matrices in which $c_{jk}=\pm 1$, thus especially the HADAMARD matrices mentioned above.

The results of the present paper have besides obvious applications in mathematical statistics also interes ing applications in number theory and the theory of sets, which will be published elsewhere.

(Received December 7, 1949.)

---

15) H. HOTELLING: Some improvements in weighing and other experimental techniques, *Annals of Math. Statistics* **15** (1944), 297—306.

16) I. e. orthogonal matrices with elements ± 1. *R. E. A. C. Paley:* On orthogonal matrices. *Journal Math. and Phys.* **12** (1933), 311 - 320.