CHARACTERIZATION OF POPULATIONS BY PROPERTIES OF SUITABLE STATISTICS

EUGENE LUKACS
OFFICE OF NAVAL RESEARCH

1. Introduction

In this paper we consider finite sets of independently and identically distributed random variables $X_1, X_2, \cdots, X_n$. Our aim is to characterize their common distribution function $F(x)$ by properties of the set $X_1, X_2, \cdots, X_n$. The problem can best be formulated by using statistical terminology.

We consider a population and a sample $X_1, X_2, \cdots, X_n$ of n independent observations drawn from this population with population distribution function $F(x)$. As usual, a measurable and single-valued function $S = S(X_1, X_2, \cdots, X_n)$ of the observations is called a statistic. Assumptions concerning the properties of the distributions of certain statistics, based on a sample from the population, will in general impose restrictions on the population distribution function $F(x)$. We are interested in assumptions which determine the population distribution function at least to the extent that it belongs to a certain family of distribution functions. Three different types of assumptions are considered.

In the first part we make assumptions which either give explicitly the distribution of S or which relate it in some specified manner to the population distribution $F(x)$.

In the second we suppose that two suitably chosen statistics $S_1 = S_1(X_1, X_2, \cdots, X_n)$ and $S_2 = S_2(X_1, X_2, \cdots, X_n)$ are given. The assumption that the statistics S_1 and S_2 are independently distributed can be used to characterize various populations. We also consider briefly the characterization of a population by the stochastic independence of more than two statistics. Finally we assume that the conditional expectation of S_1, given S_2, equals the unconditional expectation of S_1 and show that this hypothesis can also be used to characterize populations. This property is weaker than complete independence of the two statistics; its use in investigations of this kind seems to be new.

The third part deals with the characterization of populations by means of the property that two different linear statistics are identically distributed.

We denote the population distribution function by $F(x)$. The characteristic function $f(t)$ of $F(x)$ is given by

$$f(t) = \int_{-\infty}^{\infty} e^{itx} dF(x), \tag{1.1}$$

while

$$\varphi(t) = \ln f(t) \tag{1.2}$$

is called the cumulant generating function (c.g.f.) of $F(x)$. Since every characteristic function $f(t)$ is a continuous function such that $f(0) = 1$, we see that $\varphi(t)$ is certainly defined by (1.2) in an interval of the real axis which contains the origin.

A characteristic function $f(z)$ is said to be an *analytic characteristic function* if it coincides with a regular analytic function in some neighborhood of the origin in the complex z-plane.

Analytic characteristic functions were studied by several authors, [36], [37], [20]; a summary of most of their properties used in the present paper may be found in [31].

The normal distribution with mean μ and variance σ^2 will be denoted by $\Phi[(x - \mu)/\sigma]$ where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} dy . \tag{1.3}$$

Moreover, we write

$$F(x; \lambda) = \begin{cases} e^{-\lambda} \sum_{k \leq x} \lambda^k / k! & \text{for } x \geqq 0 \\ 0 & \text{for } x < 0 \end{cases} \tag{1.4}$$

for the distribution function of the Poisson distribution. The parameter λ is a real positive number. The characteristic function of (1.4) is

$$f(t; \lambda) = \exp[\lambda(e^{it} - 1)] . \tag{1.5}$$

Let α, λ be two real numbers and assume that $\lambda > 0$. The distribution

$$F(x; \alpha, \lambda) = \begin{cases} 0 & \text{for } x < 0 \\ \dfrac{\alpha^\lambda}{\Gamma(\lambda)} \displaystyle\int_0^x t^{\lambda-1} e^{-\alpha t} dt & \text{for } x > 0 \end{cases} \tag{1.6}$$

is called the Gamma distribution; its characteristic function is

$$f(t; \alpha, \lambda) = \left(1 - \frac{it}{\alpha}\right)^{-\lambda} . \tag{1.7}$$

We shall for brevity sometimes refer to the normal (Poisson, Gamma) populations and shall mean that the population distribution function $F(x)$ is the normal (Poisson, Gamma) distribution.

PART I. CHARACTERIZATIONS BY MEANS OF ONE LINEAR STATISTIC

2. Distribution function of statistic specified

In this section we consider a linear statistic

$$L = a_1 X_1 + a_2 X_2 + \cdots + a_n X_n \tag{2.1}$$

where $a_1, a_2, \cdots, a_n$ are given real numbers which are subject to the restriction

$$A_s = \sum_{j=1}^{n} (a_j)^s \neq 0 \qquad \text{for } s = 1, 2, 3, \cdots . \tag{2.2}$$

We denote the distribution function of the statistic L by $G(x)$ and write $g(t)$ and $\gamma(t)$ for its characteristic function, and its cumulant generating function, respectively. We see then that

$$\begin{aligned} g(t) &= f(a_1 t) f(a_2 t) \cdots f(a_n t) \\ \gamma(t) &= \sum_{j=1}^{n} \varphi(a_j t) . \end{aligned} \tag{2.3}$$

We assume that $g(t)$ is an analytic characteristic function so that $\gamma(t) = \ln g(t)$ is also analytic in some neighborhood of the origin. An important theorem, due to P. Lévy [20] and A. D. Raikov [36], states that the factors of analytic characteristic functions are also analytic characteristic functions. From this result and from (2.2) we conclude that $f(t)$ and $\varphi(t)$ are analytic in certain neighborhoods of the origin. It is also known that analytic characteristic functions have moments α_s, and also therefore cumulants κ_s, of any order s and that

$$\begin{aligned} \alpha_s &= E(X^s) = \int_{-\infty}^{\infty} x^s dF(x) = i^{-s} f^s(0) = i^{-s} \left.\frac{d^s f(t)}{dt^s}\right|_{t=0} \\ \kappa_s &= i^{-s}\varphi^s(0) = i^{-s} \left.\frac{d^s\varphi(t)}{dt^s}\right|_{t=0} . \end{aligned} \tag{2.4}$$

Differentiating the second equation (2.3) s-times, we obtain

$$\gamma^s(t) = \sum_{j=1}^{n} (a_j)^s \varphi^s(a_j t) . \tag{2.5}$$

We write here $\varphi^s(t)$ and $\gamma^s(t)$ for the derivative of order s of the functions $\varphi(t)$ and $\gamma(t)$.

Putting $t = 0$ in (2.5) and writing $\bar{\kappa}_s$ for the cumulant of order s of $G(x)$, we see that

$$\bar{\kappa}_s = A_s \kappa_s . \tag{2.6}$$

It follows from (2.2) and (2.6) that all the cumulants κ_s of the population distribution function $F(x)$ are determined by the cumulants $\bar{\kappa}_s$ of $G(x)$. We obtain, therefore, from (2.4) the following expansion for the c.g.f. of $F(x)$

$$\varphi(t) = \sum_{s=1}^{\infty} \frac{\bar{\kappa}_s}{A_s s!} (it)^s . \tag{2.7}$$

This is the power series expansion for the c.g.f. $\varphi(t)$. This series converges in a certain neighborhood of the origin and determines also the characteristic function $f(t)$ in some neighborhood of the origin. It is known that any analytic characteristic function is completely determined if it is given in some neighborhood of the origin. Therefore, $f(t)$, and hence also the population distribution function $F(x)$, are uniquely determined. We formulate this result in the following manner:

THEOREM 2.1. *Let $X_1, X_2, \cdots, X_n$ be a sample of n observations drawn from a population with population distribution function $F(x)$ and let $L = a_1X_1 + a_2X_2 + \cdots + a_nX_n$ be a linear statistic such that* (2.2) *is satisfied for all s. Assume that the distribution function $G(x)$ of L is given and that its characteristic function $g(t)$ is an analytic characteristic function. Then the population distribution function $F(x)$ is uniquely determined, and the relation $\kappa_s = \bar{\kappa}_s/A_s$, $s = 1, 2, \cdots$, between the cumulants $\bar{\kappa}_s$ of $G(x)$ and κ_s of $F(x)$ holds.*

Theorem 2.1 is a reformulation of results obtained by A. Rényi [38]. His interesting study of the arithmetic of distribution functions also contains other related results which could be used to characterize populations. We apply theorem 2.1 in two special cases.

First, we assume that $G(x) = \Phi[(x - \mu)/\sigma]$. Then $\bar{\kappa}_s = 0$ for $s > 2$. If all $A_s \neq 0$, we conclude from theorem 2.1 that $\varphi(t) = [\mu it/A_1] - [\sigma^2 t^2/2A_2]$, so that $F(x) = \Phi[(x - \mu A_1^{-1})\sqrt{A_2}/\sigma]$. If we take into consideration that any linear function of normally distributed random variables is normally distributed, we obtain the following characterization:

COROLLARY 2.1.1. *Let $X_1, X_2, \cdots, X_n$ be a sample from a certain population, and let*

$a_1, a_2, \cdots, a_n$ be given so that they satisfy the relations (2.2) *for all s. The distribution function of the statistic $L = a_1X_1 + \cdots + a_nX_n$ is normal with mean μ and variance σ^2 if, and only if, the population distribution function is also normal with mean μ/A_1 and variance σ^2/A_2.*

We consider next the second special case by assuming that $a_j = 1, j = 1, 2, \cdots, n$, and $G(x) = F(x; \lambda)$. Then $A_s = n$ and $\bar{\kappa}_s = \lambda$, so that according to theorem 2.1, $\varphi(t) = \lambda\,[\exp(it) - 1]/n$ and $F(x) = F(x; \lambda/n)$. Since the sum of Poisson variables obeys again Poisson's law, we can state the following:

COROLLARY 2.1.2. *Let $X_1, X_2, \cdots, X_n$ be a sample from a certain population, and denote by $\Lambda = X_1 + X_2 + \cdots + X_n$ the sum of the observations. The population has a Poisson distribution with parameter a if, and only if, the statistic Λ has a Poisson distribution with parameter na.*

Corollary 2.1.1 is a special case of a well-known theorem of H. Cramér [4], which was earlier conjectured by P. Lévy [19]. Corollary 2.1.2 is similarly implied by a theorem of D. A. Raikov [36], [35] concerning the Poisson distribution.

If we use, instead of theorem 2.1, Cramér's theorem[1] in its full generality we are able to give another characterization of the normal population. This is similar to corollary 2.1.1 but avoids the restriction (2.2) on the coefficients of the statistic L. However, in this case, neither $G(x)$ nor $F(x)$ are completely specified although their cumulants are still connected by relation (2.6).

THEOREM 2.2. *Let $X_1, X_2, \cdots, X_n$ be a sample from a certain population, and let $L = a_1X_1 + \cdots + a_nX_n$ be a linear statistic. The population is normal if, and only if, the statistic L is normally distributed.*

All the populations which we have considered so far had analytic characteristic functions. This is, however, not necessary. We conclude this from the fact that the Cauchy[2] population can be characterized by specifying the distribution of the sample mean. It is indeed easy to see that a population has the Cauchy distribution if, and only if, the mean of the sample has a Cauchy distribution.

3. A relation between the population distribution function and the distribution of S is specified

We consider a sample $X_1, X_2, \cdots, X_n$ from a population with a nondegenerate distribution function $F(x)$, and a linear statistic $L = L(a_1, a_2, \cdots, a_n) = a_1X_1 + a_2X_2 + \cdots + a_nX_n$, where $a_1, a_2, \cdots, a_n$ are arbitrary real numbers. Instead of specifying the coefficients of L and the distribution $G(x)$ of L, we make the following assumption: For *any choice* of n and of the $a_1, a_2, \cdots, a_n$, the statistic L is distributed as aX where X has the distribution function $F(x)$. Here $a = a(a_1, a_2, \cdots, a_n)$ is a real-valued function of the coefficients of L. If we denote again the distribution function of the statistic L by $G(x)$, then we can write our assumption as

$$G(x) = F\left(\frac{x}{a}\right). \tag{3.1}$$

We discuss, first, a particular case by assuming that $n = 2$, and that the coefficients a_1 and a_2 are two arbitrary positive numbers. Our assumption concerning the population

[1] The theorem and its proof may be found as theorem 19, p. 52 in [5].

[2] The c.g.f. of the Cauchy distribution is $\mu it - \lambda|t|$, $\lambda > 0$.

implies then that the distribution function $F(x)$ is stable.[3] It is known that the c.g.f. of a stable law can be written in the canonical form

$$\varphi(t) = \left(-c_0 + i c_1 \frac{t}{|t|}\right)|t|^{\alpha}, \tag{3.2}$$

where $0 < \alpha \leqq 2$, $c_0 > 0$, $|c_1 \cos(\pi\alpha/2)| \leqq c_0 \sin(\pi\alpha/2)$. The c.g.f. of any stable law satisfies for $a_1 > 0$, $a_2 > 0$, the functional equation

$$\varphi(a_1 t) + \varphi(a_2 t) = \varphi(at), \tag{3.3}$$

where a depends on a_1 and a_2.

The assumption concerning the population implies that this relation should be satisfied not only for positive, but for arbitrary a_1 and a_2. It is, therefore, necessary to determine the stable laws for which (3.3) holds for any real a_1 and a_2.

Let now a_1 and a_2 be two arbitrary real (not necessarily positive) numbers. Substituting (3.2) into (3.3), and separating real and imaginary parts, we get

$$|a_1|^{\alpha} + |a_2|^{\alpha} = |a|^{\alpha} \tag{3.4}$$

$$c_1[a_1|a_1|^{\alpha-1} + a_2|a_2|^{\alpha-1} - a|a|^{\alpha-1}] = 0. \tag{3.5}$$

We show next that the validity of (3.3) for arbitrary real a_1 and a_2 implies that $c_1 = 0$. The proof is carried indirectly; assume, therefore, that $c_1 \neq 0$ and choose a_1 and a_2 so that $a_1 a_2 < 0$.

Then $a_2|a_1|/|a_2|a_1 = -1$ and we obtain easily from (3.4) and (3.5),

$$|a_1|^{\alpha}\left(\frac{a_1}{|a_1|} - \frac{a}{|a|}\right) = |a_2|^{\alpha}\left(\frac{a_1}{|a_1|} + \frac{a}{|a|}\right). \tag{3.6}$$

An elementary discussion of this relation shows that either $a_1 = 0$ or $a_2 = 0$ in contradiction with the choice $a_1 a_2 < 0$. Therefore, $c_1 = 0$ and the c.g.f. of the population distribution is given by

$$\varphi(t) = -c_0|t|^{\alpha}, \qquad c_0 > 0,\ 0 < \alpha \leqq 2. \tag{3.7}$$

It is also seen that the dependence of a on the coefficients a_1 and a_2 is determined by (3.4).

Conversely, if the c.g.f. of a population distribution function $F(x)$ has the form (3.7) then

$$\sum_{j=1}^{n} \varphi(a_j t) = -c_0|t|^{\alpha} \sum_{j=1}^{n} |a_j|^{\alpha}. \tag{3.8}$$

Therefore, the linear form $L = a_1X_1 + a_2X_2 + \cdots + a_nX_n$ is distributed as aX, where X has the distribution $F(x)$, and where

$$|a|^{\alpha} = \sum_{j=1}^{n} |a_j|^{\alpha}. \tag{3.9}$$

We obtain, therefore, the following result.

[3] A distribution function $F(x)$ is said to be stable if to every $a_1 > 0$, $a_2 > 0$ belongs an $a > 0$ such that $F(x/a_1)*F(x/a_2) = F(x/a)$. Here the symbol * denotes the operation of convolution. This is the definition given by P. Lévy (see pp. 94 ff., pp. 198 ff. in [21]). The class of these distributions is somewhat narrower than the class which is called stable by Gnedenko and Kolmogorov (see p. 162 in [14]), which in turn was called ouasi-stable by P. Lévy (see p. 208 in [21]).

THEOREM 3.1. *Let $X_1, X_2, \cdots, X_n$ be a sample from a population with distribution function $F(x)$. Every linear statistic $L = L(a_1, a_2, \cdots, a_n) = \sum_{j=1}^{n} a_j X_j$ is distributed as $\left(\sum_{j=1}^{n} |a_j|^{\alpha}\right)^{1/\alpha} X_1$ if, and only if, $F(x)$ is a symmetric stable distribution with characteristic exponent α.*

For $\alpha = 1$, this yields a characterization of the Cauchy population; for $\alpha = 2$, of the normal population.

PART II. CHARACTERIZATION OF POPULATIONS BY THE INDEPENDENCE OF TWO STATISTICS

We consider now a sample $X_1, X_2, \cdots, X_n$ from a certain population with distribution function $F(x)$, and two statistics $S_1 = S_1(X_1, \cdots, X_n)$, $S_2 = S_2(X_1, X_2, \cdots, X_n)$. Since S_1 and S_2 are functions of the same observations, they will in general be stochastically dependent. However, for certain populations it can happen that two statistics are stochastically independent although they are functions of the same observations. In such a case one might wish to investigate whether the independence of these two statistics determines the population. The first problem of this kind which was considered concerned the normal population and will be the starting point for the following discussion.

4. Normal population characterized by the independence of one polynomial and one linear statistic

In this section we assume that the two statistics are symmetric and homogeneous polynomials in the observations. In addition, we suppose that one of the statistics is linear while the second is at least of degree two.

The best known example of this kind is the case where the statistics are the sample mean and the sample variance; in this case we obtain the following theorem.

THEOREM 4.1. *Let $X_1, X_2, X_3, \cdots, X_n$ be a sample from a certain population and denote by $\bar{X} = (X_1 + X_2 + \cdots + X_n)/n$ the sample mean, and by $M_2 = [(X_1 - \bar{X})^2 + \cdots + (X_n - \bar{X})^2]/n$ the sample variance. A necessary and sufficient condition for the normality of the population is the stochastic independence of $\bar{X}$ and M_2.*

The necessity of the condition was established by R. A. Fisher [10], while its sufficiency was investigated under various assumptions by a number of authors. R. C. Geary [12] was the first to prove the sufficiency of the condition under the restrictive assumption that the population distribution function had moments of any order. The present writer [25] gave a different proof and assumed only the existence of the second moment of the population distribution function. The assumption of the existence of moments in this and in similar problems is only used to justify the differentiations which are necessary to derive a differential equation for the characteristic function. However, the proof can be modified and the sufficiency of the condition of theorem 4.1 can be established without restrictions. This was first done by T. Kawata and H. Sakamoto [16], and somewhat later in a different manner by A. A. Zinger [41].

Theorem 4.1 can be generalized in several ways. One can, for instance, consider, instead of the sample variance, some other quadratic statistic. More generally, one can determine all populations which have the property that the sampling distribution of the mean and of a symmetric and homogeneous quadratic statistic is stochastically independent. In this case, not only the population distribution but also the quadratic statistic must be found. This was done in [26] and the following result was obtained.

THEOREM 4.2. *Let $X_1, X_2, \cdots, X_n$ be a sample from a population, and denote by $\bar{X}$ the sample mean, and by Q a homogeneous and symmetric quadratic statistic. Assume that the second moment of the population distribution function exists. The statistic Q is stochastically independent of $\bar{X}$ if, and only if, one of the following three mutually exclusive conditions is satisfied.*

(a) *The population distribution function is a (nondegenerate) normal distribution and Q is proportional to the variance.*

(b) *The population distribution is degenerate.*[4]

(c) *The population distribution function is a step function with two symmetrically located steps and $Q = X_1^2 + X_2^2 + \cdots + X_n^2$.*

This theorem somewhat exceeds the framework of this section since it characterizes in case (c) also a nonnormal population.

It should not be difficult to give further generalizations, for example, by considering inhomogeneous quadratic statistics, or by avoiding the assumption that the second moment of the population distribution function exists.

Another generalization of theorem 4.1 is obtained if instead of the variance one considers symmetric and homogeneous polynomial statistics of higher degree. We select for this statistic the k-statistic of order p.

DEFINITION 4.1. *The k-statistic of order p is a symmetric and homogeneous polynomial statistic of degree p, such that $E(k_p) = \kappa_p$ for any population distribution which has moments up to the order p.*

THEOREM 4.3. *Let $X_1, X_2, \cdots, X_n$ be a sample of n observations taken from a population with population distribution function $F(x)$, and denote by p an integer greater than one. Assume that the pth moment of $F(x)$ exists. The population is normal if, and only if, the k-statistic of order p is independent of the sample mean.*

Theorem 4.3 was independently proved by this writer [28], and by D. Basu and R. G. Laha [1]. This theorem generalizes theorem 4.1 by using the polynomial statistic k_p of order p, instead of the sample variance M_2. A more obvious approach would have been to consider the independence of the sample mean $\bar{X}$, and the sample moment of order p about the mean. While it seems likely that this approach could be used to characterize the normal population, the discussion becomes rather awkward, even in the simplest case $p = 4$.

We indicate briefly the proof of theorem 4.3. The independence of k_p and k_1 can be written as

$$E\,[\exp\,(itk_1 + iuk_p)\,] = E\,(e^{itk_1})\,E\,(e^{iuk_p})\,. \tag{4.1}$$

Since, according to the assumptions of the theorem, the pth moment of $F(x)$ exists, one can differentiate (4.1) with respect to u, and put afterwards $u = 0$. This yields

$$E\,(k_p e^{itnk_1}) = \kappa_p\,[\,f\,(t)\,]^{\,n}. \tag{4.2}$$

We have seen in section 1 that the c.g.f. $\varphi(t)$ of a distribution function, $F(x)$, is always defined in some neighborhood $-\Delta < t < \Delta$ of the origin. If we restrict ourselves to this neighborhood we can easily compute the left-hand side of (4.2) by means of Faà di Bruno's formula [28], [9], and obtain

$$E\,(k_p e^{itnk_1}) = i^{-p}\,[\,f\,(t)\,]^{\,n}\,\frac{d^p\varphi\,(t)}{dt^p} \tag{4.3}$$

[4] In this case no restriction is imposed on Q.

so that

$$\frac{d^p\varphi(t)}{dt^p} = i^p\kappa_p \qquad \text{for } |t| < \Delta. \tag{4.4}$$

This is a differential equation for the c.g.f. Using the initial conditions $\varphi^j(0) = i^j\kappa_j$, $j = 1, 2, \cdots, (p - 1)$, we obtain the solution

$$\varphi(t) = \sum_{j=1}^{p} \frac{\kappa_j (it)^j}{j!}. \tag{4.5}$$

Since equation (4.4) is valid only for $|t| < \Delta$, the solution obtained in this manner is valid so far only for $|t| < \Delta$. Its range of validity can, however, be extended to the full real axis by applying the following lemma.

LEMMA 4.1. *Let $A(z)$ be a function of the complex variable z, which is regular in some neighborhood $|z| < \rho$ of the origin. Let $f(t)$ be a characteristic function, and $\Delta > 0$ an arbitrary positive number, and assume that $f(t) = A(t)$ if t is real, and if $-\Delta < t < \Delta$. Then $f(t)$ is an analytic characteristic function.*

This lemma was found by R. P. Boas [3]; more recently a special case was rediscovered by Yu. V. Linnik (see theorem 4 in [23]). Lemma 4.1 is very useful in work of this kind, and can often be employed instead of continuity considerations, which some authors used to extend the validity of the solution to all real values of t. In some instances (see, for example, [17], [25], [26]) the discussion of this extension was not given. A proof of lemma 4.1 is given in the appendix. We see, therefore, from (4.5) and from lemma 4.1, that

$$f(t) = \exp\left[\sum_{j=1}^{p} \frac{\kappa_j (it)^j}{j!}\right]. \tag{4.6}$$

In order to find the distribution function whose characteristic function satisfies (4.4), we must select among the functions (4.6) those which are characteristic functions. This is done easily by means of the following result [32]:

FIRST THEOREM OF MARCINKIEWICZ.[5] *No function of the form* (4.6) *with $p > 2$, $\kappa_p \neq 0$ can be a characteristic function.*

But this means that $F(x)$ is a normal distribution and proves the sufficiency of the condition of theorem 4.3. The necessity of the condition follows immediately from the translation-invariance of the k-statistics and from a result due to J. F. Daly [6], who showed that in a normal population every translation-invariant statistic is stochastically independent from the sample mean.

This fact suggests the possibility of characterizing a normal population by the independence of the sample mean and some translation-invariant statistic, other than a k-statistic. One could ask for the conditions which a translation-invariant statistic, or a system of such statistics, must satisfy, in order that its independence from the sample mean should imply the normality of the population.

The following result was obtained by this writer.

THEOREM 4.4. *Let $S_r(x_1, x_2, \cdots, x_n)$, $r = 1, 2, \cdots, (n - 1)$, be $(n - 1)$ single-valued and measurable functions such that*

(i) *$S_r(x_1 + a, x_2 + a, \cdots, x_n + a) = S_r(x_1, x_2, \cdots, x_n)$, $r = 1, 2, \cdots, (n - 1)$, for any real a. Suppose that the system of equations*

[5] Actually Marcinkiewicz proved a more general theorem: No entire function of finite order $p > 2$ whose exponent of convergence $q < p$ can be a characteristic function.

(ii) $S_r(y_1, y_2, \cdots, y_{n-1}, 0) = z_r$, $r = 1, 2, \cdots, (n-1)$, *has a solution*

(iii) $y_r = h_r(z_1, z_2, \cdots, z_{n-1})$, $r = 1, 2, \cdots, (n-1)$, *such that the* $h_r(z_1, z_2, \cdots, z_{n-1})$ *are single-valued and measurable functions. Let* $X_1, X_2, \cdots, X_n$ *be a sample taken from a certain population. The population is normal if, and only if, the translation-invariant statistics* $S_1(X_1, X_2, \cdots, X_n), S_2(X_1, X_2, \cdots, X_n), \cdots, S_{n-1}(X_1, X_2, \cdots, X_n)$ *are independently distributed from the sample mean* $\bar{X} = (X_1 + X_2 + \cdots + X_n)/n$.

The proof uses a particular case of theorem 5.1 which will be discussed in the next section. We remark that it is not necessary to assume the existence of moments of the population distribution function.

5. Normal population characterized by the independence of two linear statistics

The problem discussed in this section is the determination of the population for which two linear statistics are independently distributed. The problem was attacked independently by several authors, [15], [2], [11], who treated the case of a sample of two, mostly under various restrictive assumptions (such as existence of a frequency function, existence of second moments). B. V. Gnedenko [13] and, independently, G. Darmois [7] used finite differences to avoid differentiation and the need for the assumption of the existence of moments. The general case, as formulated below in theorem 5.1, can be treated without any restrictive assumptions. This theorem was first obtained in full generality by V. P. Skitovich [39]. A detailed proof, and some general results concerning systems of linear forms, may be found in a more recent paper [40] of the same author.

THEOREM 5.1. *Let* $X_1, X_2, \cdots, X_n$ *be a sample from a certain population and consider the linear forms,* $L_1 = a_1X_1 + a_2X_2 + \cdots + a_nX_n$ *and* $L_2 = b_1X_1 + b_2X_2 + \cdots + b_nX_n$, *such that* $\sum a_sb_s = 0$ *while* $\sum (a_sb_s)^2 \neq 0$. *The population is normal if, and only if,* L_1 *and* L_2 *are independently distributed.*

The theorem was also proved [30] by means of the method used in section 4; however, this approach required the existence of the nth moment of the population distribution. A proof of theorem 5.1 is also indicated in [8] where connections with problems of factor analysis are also discussed. Further references may be found in [30] and in [8].

6. Characterization by constant regression

If one examines the various papers, [12], [25], [16], [40], treating theorem 4.1, one notices that gradually all unessential assumptions were removed, and that finally only the independence of the two statistics was supposed. However, a close scrutiny of the proofs reveals that none of them uses the assumption of the stochastic independence of the sample mean and of the sample variance fully. The same remark applies also to the proofs of theorems 4.2 and 4.3 (but not to the proof of theorem 4.4 which has a different structure).

This situation leads very naturally to the question whether the assumptions of theorem 4.1 could not be weakened, or at least be modified, to such an extent that the assumptions are fully used.

The discussion of this approach calls for the introduction of an appropriate terminology.

Let X and Y be two random variables, and assume that the first moment of Y exists. We denote, as usual, by $E(Y)$ the expected value of Y, and write $E(Y|X)$ for the conditional expectation of Y, given the value of X. We introduce the following definition:

DEFINITION 6.1. *A random variable Y, which has finite expectation, is said to have constant regression on a random variable X, if the relation*

$$E(Y \mid X) = E(Y) \tag{6.1}$$

holds almost everywhere.

We need the following lemma:

LEMMA 6.1. *Let X and Y be two random variables and assume that the expecion $E(Y)$ exists. The random variable Y has constant regression on X if, and only if, the relation*

$$E(Y e^{itX}) = E(Y) E(e^{itX}) \tag{6.2}$$

holds for all real t.

If one multiplies relation (6.1) by $\exp(itX)$ and then takes the expectation, one sees immediately that the condition is necessary.

To prove its sufficiency we assume (6.2), and we must distinguish two cases. First, we suppose that

$$E(Y) \neq 0 . \tag{6.3}$$

We can then write (6.2) as

$$E\left[\frac{e^{itX} E(Y \mid X)}{E(Y)}\right] = E(e^{itX}) . \tag{6.4}$$

The functions under the first expectation sign depend only on X; it is therefore possible to take these expectations with respect to the marginal distribution P of the random variable X. We rewrite (6.4) therefore in the form

$$\int_{R_1} \frac{e^{itx} E(Y \mid x)}{E(Y)} dP = \int_{R_1} e^{itx} dP \tag{6.5}$$

where we integrate over the one-dimensional space R_1 of the random variable X. We introduce the set function

$$\nu(A) = \int_A \frac{E(Y \mid x)}{E(Y)} dP \tag{6.6}$$

which is defined on the Borel sets of R_1. This is a function of bounded variation and we see from (6.5) and (6.6) that

$$\int_{R_1} e^{itx} d\nu = \int_{R_1} e^{itx} dP . \tag{6.7}$$

Since the uniqueness theorem for characteristic functions is valid for the Fourier transforms of functions of bounded variation, we conclude that $\nu(A) = P(A)$ or

$$\int_A \frac{E(Y \mid x)}{E(Y)} dP = \int_A dP . \tag{6.8}$$

As P is obviously absolutely continuous with respect to P, we must have $E(Y \mid X)/E(Y) = 1$ almost everywhere; but this means that Y has constant regression on X. We still must consider the case where

$$E(Y) = 0 . \tag{6.3a}$$

In this case, (6.2) reduces to $E[Y \exp(itX)] = E[\exp(itX) E(Y \mid X)] = 0$, or written as an integral

$$\int_{R_1} e^{itx} E(Y \mid X) dP = 0 . \tag{6.9}$$

If we introduce the set function $\mu(A)$ defined on R_1 by $\mu(A) = \int_A E(Y|X)dP$, we can write (6.9) as $\int_{R_1} \exp(itx)d\mu = 0$. From this we conclude that $\mu(A)$ is constant so that $\mu(A) = \mu(R_1) = E(Y) = 0$. But this means that $E(Y|X) = 0$ almost everywhere. We see then from (6.3a) that Y has constant regression on X. This completes the proof of the lemma.

We use next lemma 6.1 to obtain a characterization of the normal population related to the statement (a) of theorem 4.2.

THEOREM 6.1. *Let $X_1, X_2, \cdots, X_n$ be a sample from a nondegenerate population, and assume that the second moment of the population distribution exists. Denote by $\Lambda = X_1 + X_2 + \cdots + X_n$, the sum of the observations, and by $Q = \sum_{r=1}^{n}\sum_{s=1}^{n} a_{rs}X_rX_s$, a quadratic statistic such that*

$$B_1 = \sum_{r=1}^{n} a_{rr} \neq 0\,, \qquad B_2 = \sum_{r=1}^{n}\sum_{s=1}^{n} a_{rs} = 0\,. \tag{6.10}$$

The population is normal if, and only if, the statistic Q has constant regression on Λ.

We first prove the sufficiency of the condition. It follows immediately from lemma 6.1 that $E[Q \exp(it\Lambda)] = E(Q)[f(t)]^n$, where $f(t)$ is the characteristic function of the population distribution function. A simple computation yields then $B_1f''(f)^{n-1} - B_1(f')^2(f)^{n-2} = -(f)^nE(Q) = -B_1\kappa_2(f)^n$, where we write, for simplicity, f, f', f'', instead of $f(t), f'(t), f''(t)$. We first restrict the variable t to an interval around the origin in which $f(t)$ has no zeros. We obtain then for the c.g.f. $\varphi(t)$ the equation $\varphi''(t) = -\kappa_2$ so that $\varphi(t) = -\kappa_2t^2/2 + i\kappa_1t$. This solution was obtained only for a certain neighborhood of the origin, but can be extended to all real t according to lemma 4.1.

To prove the necessity of the condition we observe that the c.g.f. of the normal distribution satisfies the differential equation $B_1\varphi'' = -B_1\kappa_2$ or $B_1f''(f)^{n-1} - B_1(f')^2(f)^{n-2} = -E(Q)[f]^n$. Using (6.10) this can be written as $E[Q \exp(it\Lambda)] = E(Q)\, E[\exp(it\Lambda)]$. The necessity of the condition follows from this equation and from lemma 6.1. The sufficiency of the condition of theorem 6.1 was recently proved by R. G. Laha [17]. Condition (6.10) means that the expected value of the statistic Q is proportional to the variance, in other words, that Q is proportional to an unbiased estimate of the population variance.

We consider next another quadratic statistic. Here (6.10) is replaced by a different requirement.

THEOREM 6.2. *Let $X_1, X_2, \cdots, X_n$ be a sample from a nondegenerate population, and assume that the second moment of the population distribution function exists.*

Denote by $Q = \sum_{r=1}^{n}\sum_{s=1}^{n} a_{rs}X_rX_s$ a quadratic statistic such that

$$E(Q) = 0 \tag{6.11}$$

and also

$$B_1 = \sum_{r=1}^{n} a_{rr} \neq 0\,, \qquad B_2 = \sum_{r=1}^{n}\sum_{s=1}^{n} a_{rs} \neq 0\,. \tag{6.12}$$

The population is a Gamma population if, and only if, the statistic Q has constant regression on $\Lambda = X_1 + X_2 + \cdots + X_n$.

We first prove the sufficiency of the condition. From lemma 6.1 and equation (6.11) it follows that $E[Q \exp(it\Lambda)] = 0$. In the same manner as before we transform this into a differential equation

$$B_1\varphi'' + B_2(\varphi')^2 = 0 \tag{6.13}$$

for the c.g.f. $\varphi(t)$. Equation (6.13) is valid in an interval $|t| < \Delta$. From (6.11) it is seen that $B_1\kappa_2 + B_2\kappa_1^2 = 0$; we introduce $\lambda = -B_1/B_2 = \kappa_1^2/\kappa_2 > 0$, and find easily the solution

$$f(t) = \left(1 - \frac{i\kappa_1 t}{\lambda}\right)^{-\lambda}. \tag{6.14}$$

This solution is originally obtained for $|t| < \Delta$; its validity for all real t follows from lemma 4.1. Clearly, (6.14) is the characteristic function of the well-known Gamma distribution.

The necessity of the condition follows again almost immediately from lemma 6.1, and the fact that the c.g.f. of a Gamma distribution satisfies the differential equation $\varphi'' = (\varphi')^2/\lambda$.

The next theorem describes the Poisson population, that is, a population whose distribution function is given by (1.4). We use again the notation of section 4 and write k_s for the k-statistic of order s. It is also convenient to use the following terminology.

DEFINITION 6.2. *A (finite) point a is said to be the left extremity of a distribution $F(x)$, if $F(x) = 0$ for $x < a$, while $F(x) > 0$ for $x > a$. We write then $a =$ lext $[F]$.*

In a similar manner one could define the right extremity of a distribution.

DEFINITION 6.3. *A distribution is said to be one-sided if it has one extremity; finite, if it has both extremities.*

G. Pólya [34] has derived formulas which express the extremities of a distribution $F(x)$ in terms of its characteristic function $f(t)$. Pólya formulated his results for finite distributions. It is, however, easy to see that his formulas for the extremities apply also to one-sided distributions with analytic c.f. We shall need the formula, for the left extremity, namely,

$$\text{lext}\,[F] = -\overline{\lim_{r\to\infty}}\, r^{-1} \log|f(ir)|. \tag{6.15}$$

THEOREM 6.3. *Let $X_1, X_2, \cdots, X_n$ be a sample from a population with population distribution function $F(x)$ and denote by $p \geqq 1$ a positive integer. Assume that the $(p+2)$nd moment of $F(x)$ exists and suppose also that $F(x)$ is a one-sided distribution which has the point $x = 0$ as its left extremity. The population distribution function is the Poisson distribution* (1.4) *if, and only if, the statistic $S = k_{p+2} - k_p$ has constant regression on k_1.*

We first prove the sufficiency of the condition and assume that $k_{p+2} - k_p$ has constant regression on k_1; clearly $k_{p+2} - k_p$ also has constant regression on nk_1 and according to lemma 6.1 we can write

$$E\{(k_{p+2} - k_p)\, e^{itnk_1}\} = (\kappa_{p+2} - \kappa_p)[f(t)]^n. \tag{6.16}$$

This equation can be simplified by expressing the left-hand side in terms of the c.g.f. $\varphi(t)$, and we obtain

$$\frac{d^{p+2}\varphi(t)}{dt^{p+2}} + \frac{d^p\varphi(t)}{dt^p} = i^{p+2}(\kappa_{p+2} - \kappa_p). \tag{6.17}$$

This relation is again valid in some interval $|t| < \Delta$. Equation (6.17) is an inhomogeneous linear differential equation with constant coefficients. Its solution is

$$\varphi(t) = c_1 e^{it} + c_2 e^{-it} + d_0 + \sum_{s=1}^{p-1} \frac{d_s}{s!} (it)^s - (\kappa_{p+2} - \kappa_p) \frac{(it)^p}{p!}. \tag{6.18}$$

The coefficients c_1, c_2, d_0, $d_1, \cdots, d_{p-1}$ can be determined from the initial conditions $\varphi(0) = 0$, $\varphi^s(0) = i^s \kappa_s$, $s = 1, 2, \cdots, (p+1)$, and we find

$$\begin{gathered} c_1 = \frac{\kappa_{p+2} + \kappa_{p+1}}{2}, \qquad c_2 = (-1)^p \frac{\kappa_{p+2} - \kappa_p}{2}, \\ d_s = \kappa_s - c_1 - (-1)^s c_2 \qquad \text{for} \qquad s = 0, 1, 2, \cdots, (p-1). \end{gathered} \tag{6.19}$$

We conclude from lemma 4.1 that the solution (6.18) is valid for all real t. We use now the assumption that $\operatorname{lext}[F] = 0$ in order to determine $\varphi(t)$ completely. It is easily seen from (6.15) and (6.18) that $\operatorname{lext}[F] = 0$ implies that

$$c_2 = 0, \kappa_{p+2} = \kappa_p, \qquad d_0 = -c_1, \ d_s = 0 \text{ for } s = 1, 2, \cdots, (p-1). \tag{6.20}$$

But then (6.18) reduces to

$$\varphi(t) = c_1 (e^{it} - 1). \tag{6.21}$$

This is the c.g.f. of the Poisson distribution.

The proof of the necessity of the condition is quite analogous to the corresponding proofs for theorems 6.1 and 6.2. A special case of theorem 6.3 was recently proved by this writer [29].

7. The Gamma population

In theorem 6.2 the Gamma population was characterized by the fact that a certain quadratic statistic Q has constant regression on the sample mean $\bar{X}$. It can easily be shown that this particular statistic Q cannot be stochastically independent of $\bar{X}$ in a Gamma population. Therefore, the use of the weaker assumption of constant regression is quite essential in this theorem. In the present section we turn to the characterization of a Gamma population by two stochastically independent statistics.

An interesting property of Gamma populations was proved by E. J. G. Pitman [33]: If $X_1, X_2, \cdots, X_n$ is a sample drawn from a Gamma population, then the sum $\Lambda = X_1 + X_2 + \cdots + X_n$ is always distributed independently of any scale-invariant statistic. A statistic $S(X_1, \cdots, X_n)$ is said to be scale invariant if $S(aX_1, aX_2, \cdots, aX_n) = S(X_1, X_2, \cdots, X_n)$ for any real $a \neq 0$.

This property of the Gamma population suggests the question whether it is possible to characterize the Gamma population by the stochastic independence of a scale-invariant statistic and the sample mean. The answer is in the affirmative; an example is given by theorem 7.1.

THEOREM 7.1. *Let $X_1, X_2, \cdots, X_n$ be a sample from a nondegenerate population and assume that the second moment of the population distribution function exists. Denote by $\Lambda = X_1 + X_2 + \cdots + X_n$ a linear statistic proportional to the sample mean, and by $Q = \sum_{r=1}^{n} \sum_{s=1}^{n} a_{rs} X_r X_s$ a quadratic statistic with real coefficients a_{rs}. We write $B_1 = \sum_{r=1}^{n} a_{rr}$,*

$B_2 = \sum_{r=1}^{n} \sum_{s=1}^{n} a_{rs}$, *and suppose that* $B_2 \neq nB_1$. *The population is a Gamma population if, and only if, the statistic* $S = Q/\Lambda^2$ *has constant regression on* Λ.

Theorem 7.1 was found by R. G. Laha [18]. It can be proved by the method used for the theorems of the preceding section. It admits an interesting corollary which is also due to Laha.

COROLLARY 7.1.1. *Let* $X_1, X_2, \cdots, X_n$ *be a sample from a nondegenerate population and assume that the second moment of the population distribution function exists. The population is the Gamma population if, and only if, the statistics* $S = (a_1X_1 + a_2X_2 + \cdots + a_nX_n)/\Lambda$, *and* $\Lambda = X_1 + X_2 + \cdots + X_n$ *are independently distributed.*

If S and Λ are independently distributed, then the same is true for S^2 and Λ, and therefore S^2 has *a fortiori* constant regression on Λ. The corollary follows then from theorem 7.1.

In formulating theorem 7.1 it would have been permissible to replace the requirement that S has constant regression on Λ by the assumption that Λ and S are stochastically independent. The weaker assumption is, however, sufficient to establish the theorem.

This situation suggests the question whether it is possible to assume in theorems of this type full independence, but not to require the existence of moments. The present writer has recently obtained a result in this direction. This concerns the Gamma distribution rather than the Gamma population but could probably be modified to characterize the population.

THEOREM 7.2. *Let* X *and* Y *be two nondegenerate and positive random variables, and assume that they are independently distributed. The random variables* $U = X + Y$ *and* $V = X/Y$ *are independent if, and only if, both* X *and* Y *have Gamma distributions with the same scale parameter.*

It should be noted that it is not required that X and Y be identically distributed. We finally characterize the Gamma population by the independence of a system of scale-invariant statistics from the sample mean.

THEOREM 7.3. *Let* $S_r(x_1, x_2, \cdots, x_n)$, $r = 1, 2, \cdots, (n-1)$, *be* $(n-1)$ *single-valued and measurable functions such that*

(i) $S_r(ax_1, ax_2, \cdots, ax_n) = S_r(x_1, x_2, \cdots, x_n)$, $r = 1, 2, \cdots, (n-1)$, *for any real* $a \neq 0$. *Suppose that the system of equations*

(ii) $S_r(y_1, y_2, \cdots, y_{n-1}, 1) = z_r$, $r = 1, 2, \cdots, (n-1)$, *has a solution*

(iii) $y_r = h_r(z_1, z_2, \cdots, z_{n-1})$, $r = 1, 2, \cdots, (n-1)$, *such that the* $h_r(z_1, z_2, \cdots, z_{n-1})$ *are single-valued and measurable functions. Let* $X_1, X_2, \cdots, X_n$ *be a sample taken from a population with population distribution function* $F(x)$ *and assume that* $F(+0) = 0$ *and that* lext $[F] = 0$. *The population is a Gamma population if, and only if, the scale-invariant statistics* $S_1(X_1, X_2, \cdots, X_n)$, $S_2(X_1, X_2, \cdots, X_n), \cdots, S_{n-1}(X_1, X_2, \cdots, X_n)$ *are stochastically independent from the sample mean* $\bar{X} = (X_1 + X_2 + \cdots + X_n)/n$.

The necessity of the condition follows from Pitman's result so that we have to prove only its sufficiency. We denote by

$$\theta_r = S_r\left(\frac{X_1}{X_n}, \frac{X_2}{X_n}, \cdots, \frac{X_{n-1}}{X_n}, 1\right), \quad r = 1, 2, \cdots, (n-1). \tag{7.1}$$

From the assumption that the statistics $S_r(X_1, X_2, \cdots, X_n)$, $r = 1, 2, \cdots, (n-1)$, are independent of $\bar{X}$ and from (i) we conclude that also the statistics $\theta_1, \theta_2, \cdots, \theta_{n-1}$ are independent of $\bar{X}$. Hence, any single-valued and measurable function of the $\theta_1, \theta_2, \cdots,$

θ_{n-1} will also be independent of $\bar{X}$, or equivalently, of $n\bar{X} = (X_1 + X_2 + \cdots + X_n)$. In particular, the system

$$(7.2) \qquad h_r(\theta_1, \theta_2, \cdots, \theta_{n-1}) \qquad r = 1, 2, \cdots, (n-1),$$

is stochastically independent of $(X_1 + X_2 + \cdots + X_n)$. From (ii) and (iii) it follows that

$$(7.3) \qquad h_r(\theta_1, \theta_2, \cdots, \theta_{n-1}) = \frac{X_r}{X_n}, \qquad r = 1, 2, \cdots, (n-1).$$

Therefore, the random variables (X_1/X_n), $(X_2/X_n), \cdots, (X_{n-1}/X_n)$ are independent of $(X_1 + X_2 + \cdots + X_n)$. If we express this in terms of characteristic functions we obtain

$$(7.4) \quad E\left\{\exp\left[i\sum_{r=1}^{n-1}\frac{u_r X_r}{X_n} + it\sum_{r=1}^{n} X_r\right]\right\} = E\left\{\exp\left[i\sum_{r=1}^{n-1}\frac{u_r X_r}{X_n}\right]\right\}E\left[\exp\left(it\sum_{r=1}^{n} X_r\right)\right].$$

Let k be an integer such that $1 \leqq k < n$, and put in (7.4) $u_r = 0$; if $r \neq k$, then it is easily seen that

$$(7.5) \quad E\left\{\exp\left[\frac{iu_k X_k}{X_n} + it(X_k + X_n)\right]\right\} = E\left\{\exp\frac{iu_k X_k}{X_n}\right\}E\{\exp[it(X_k + X_n)]\}.$$

This means that the random variables (X_k/X_n) and $(X_k + X_n)$ are independent; the statement of theorem 7.3 follows then immediately from theorem 7.2.

PART III. CHARACTERIZATION OF POPULATIONS BY IDENTICALLY DISTRIBUTED STATISTICS

We consider in the following two different statistics, $S_1 = S_1(X_1, X_2, \cdots, X_n)$ and $S_2 = S_2(X_1, X_2, \cdots, X_n)$. In general they have different distributions; however, it can happen that for certain populations two different statistics are identically distributed. A simple example is the normal population where two different linear statistics, $L_1 = a_1X_1 + a_2X_2 + \cdots + a_nX_n$ and $L_2 = b_1X_1 + b_2X_2 + \cdots + b_nX_n$, are identically distributed provided that $\sum_{j=1}^{n} a_j = \sum_{j=1}^{n} b_j$ and $\sum_{j=1}^{n} a_j^2 = \sum_{j=1}^{n} b_j^2$. The possibility exists, therefore, that a population, which admits two different but identically distributed statistics, may be characterized by this property.

8. Linear statistics and the normal population

The first study of this kind was undertaken by J. Marcinkiewicz [32] who obtained the following result:

SECOND THEOREM OF MARCINKIEWICZ. *Let $X_1, X_2, \cdots$ be a finite or infinite sequence of identically distributed random variables with common distribution function $F(x)$. Suppose that $F(x)$ has moments of any order, and assume that the two (finite or infinite) sums, $\sum a_jX_j$ and $\sum b_jX_j$, exist and are identically distributed. Then either the sequences $\{|a_j|\}$ and $\{|b_j|\}$ are identical, except for the order of the terms, or the distribution $F(x)$ is normal (possibly degenerate).*

This theorem gives a sufficient condition for the normality of $F(x)$. In statistical investigations we use only finite samples and therefore we are here interested only in the case of finite sums. This case is much simpler than the case of infinite sums. We formulate it now as a characterization of the normal population.

THEOREM 8.1. *Let $X_1, X_2, \cdots, X_n$ be a sample from a population with population distribution function $F(x)$, and assume that all absolute moments*

$$\beta_k = \int_{-\infty}^{\infty} |x|^k dF(x), \qquad k = 1, 2, 3, \cdots \tag{8.1}$$

exist. Consider two linear statistics

$$L_1 = \sum_{r=1}^{n} a_r X_r \quad \text{and} \quad L_2 = \sum_{r=1}^{n} b_r X_r, \tag{8.2}$$

and suppose that the numbers $|a_1|, |a_2|, \cdots, |a_n|$ are not a permutation of the numbers $|b_1|, |b_2|, \cdots, |b_n|$ and that

$$\sum_{r=1}^{n} a_r = \sum_{r=1}^{n} b_r, \qquad \sum_{r=1}^{n} a_r^2 = \sum_{r=1}^{n} b_r^2. \tag{8.3}$$

The necessary and sufficient condition for the normality of the population is that the statistics L_1 and L_2 be identically distributed.

We indicate briefly the proof of this theorem and follow closely the method used by Marcinkiewicz. The necessity of the condition is deduced from elementary properties of the normal distribution. To prove the sufficiency, we assume that our population has the property that the two statistics (8.2) are identically distributed. Then the population with population distribution function $G(x) = F(x)*[1 - F(-x)]$ also has this property. In terms of the c.g.f. $\gamma(t)$ of $G(x)$, this means that

$$\sum_{s=1}^{n} \gamma(a_s t) = \sum_{s=1}^{n} \gamma(b_s t) \tag{8.4}$$

in a certain neighborhood of the origin.

Since all the moments of $G(x)$ exist, we may differentiate (8.4) any number of times. Let k be a positive integer; we differentiate (8.4) $2k$-times and put then $t = 0$; in this manner, we get

$$\left[\sum_{s=1}^{n} (a_s)^{2k} - \sum_{s=1}^{n} (b_s)^{2k}\right] \gamma^{2k}(0) = 0. \tag{8.5}$$

It is easily seen that the relation

$$\sum_{s=1}^{n} (a_s)^{2k} = \sum_{s=1}^{n} (b_s)^{2k} \tag{8.6}$$

can hold for infinitely many k only if the $|a_1|, |a_2|, \cdots, |a_n|$ are permutations of the $|b_1|, |b_2|, \cdots, |b_n|$. Therefore, the relation $\gamma^{2k}(0) = 0$ must hold for almost all k. Since $G(x)$ is a symmetric distribution, we have $\gamma^{2k-1}(0) = 0$ for all k, so that there exists an integer p such that $\gamma^k(0) = 0$ for $k > p$. This means, however, that the c.g.f. of $G(x)$ is a polynomial of degree not exceeding p; hence, by the first theorem of Marcinkiewicz, $G(x)$ is normal. It follows then from Cramér's theorem [4] that $F(x)$ is also normal.

Recently Yu. V. Linnik considerably generalized Marcinkiewicz's work. He published several papers [22], [23], [24] dealing with the determination of populations for which two given statistics are identically distributed. He obtained a necessary and sufficient condition for the equivalence of the statement that the population is normal with the assertion that two linear statistics are identically distributed. Furthermore, Linnik derived sufficient conditions for the normality of a population which admits two identically distributed linear statistics. He also characterized a class of symmetric distributions which contains the convolutions of symmetric stable laws. While most of this work concerns linear statistics, [22] contains indications that certain nonanalytic statistics can be used in a similar manner.

APPENDIX

We give here the proof of lemma 4.1.

Let R be a positive number such that $R < \min(\Delta, \rho)$, and denote by C the circle $|z| = R$. According to the assumptions of lemma 4.1, the nth derivative $f^n(0)$ of the characteristic function exists at the origin, and can be found by means of Cauchy's formula as

$$f^n(0) = \frac{n!}{2\pi i} \int_C \frac{A(z)\,dz}{z^{n+1}}. \tag{A.1}$$

We denote by α_n (β_n respectively) the nth moment (nth absolute moment) of the distribution $F(x)$, corresponding to $f(t)$, and write $M(R) = \max\limits_{|z| \leq R} |A(z)|$. Then

$$|\alpha_n| = \frac{n!}{2\pi} \left| \int_C \frac{A(z)\,dz}{z^{n+1}} \right| \leq \frac{n!M(R)}{R^n} \tag{A.2}$$

for all n. Using the inequality $2\beta_{2n-1} \leqq \alpha_{2n} + \alpha_{2n-2}$, we get upper bounds for the absolute moments. Let r be a real number such that $|r| < R$; using these bounds it is easy to show that the series

$$S = \sum_{n=0}^{\infty} \frac{\beta_n r^n}{n!} \tag{A.3}$$

has a convergent majorant and is therefore convergent for $|r| < R$. Let $a > 0$ and $b > 0$ be two arbitrary positive numbers; then

$$S = \sum_{n=0}^{\infty} \frac{1}{n!} \left[\int_{-\infty}^{\infty} |xr|^n dF(x) \right] \geqq \sum_{n=0}^{\infty} \int_{-a}^{b} \frac{|xr|^n}{n!} dF(x). \tag{A.4}$$

Since

$$\int_{-a}^{b} e^{|xr|} dF(x) = \sum_{n=0}^{\infty} \int_{-a}^{b} \frac{|xr|^n}{n!} dF(x), \tag{A.5}$$

we see from (9.4) that

$$S \geqq \int_{-a}^{b} e^{|xr|} dF(x) \geqq \int_{-a}^{b} e^{xr} dF(x) \tag{A.6}$$

for any finite a and b. This implies the existence of the integral $\int_{-\infty}^{\infty} \exp(rx)dF(x)$ for $|r| < R$. From this fact one can conclude that $f(t)$ is an analytic characteristic function.

ADDENDUM
(Added October 31, 1955)

The author obtained through the courtesy of Messrs. C. R. Rao and R. G. Laha notes of lectures held by Professor Yu. V. Linnik in Calcutta in the fall of 1954. These notes [42] indicate that mathematicians at Leningrad University are actively engaged in research leading to the characterization of populations either by the independence or by the equidistribution of statistics. Some of this work [22], [23], [24], [39], [40] was discussed in the main body of this paper; the purpose of this addendum is to list two more recent results.

The first of these relates the characterization of a population by independent statistics (discussed in part II) to the characterization by equidistributed statistics (treated in part III). The second result deals with the possibility of characterizing the normal population by the independence of certain translation-invariant statistics from the sample mean.

We denote in the following by $\vec{X} = (X_1, X_2, \cdots, X_n)$ a sample of n independent observations and write for brevity $S(\vec{X}) = S(X_1, X_2, \cdots, X_n)$ for a statistic based on this sample. Then the following theorem holds.

THEOREM A1. *Let $\vec{X}$ and $\vec{Y}$ be two samples of size n taken from an arbitrary population and denote by S_1 and S_2 two statistics. The statistics $S_1(\vec{X})$ and $S_2(\vec{X})$ are independently distributed if, and only if, the two statistics $t_1S_1(\vec{X}) + t_2S_2(\vec{X})$ and $t_1S_1(\vec{X}) + t_2S_2(\vec{Y})$ are identically distributed for all real t_1 and t_2.*

The theorem is proved by writing the condition for equidistribution, respectively, for independence in terms of characteristic functions. Theorem A1 is a slightly specialized version of the result given in [42].

At the end of section 4 we mentioned the fact that in a normal population every translation-invariant statistic is independent of the sample mean and we raised the question what conditions a translation-invariant statistic, or a system of such statistics, must satisfy in order that its independence from the sample mean should imply the normality of the population. A partial answer, referring to systems of translation-invariant statistics, was given by the theorem 4.4; a different approach, yielding also only a partial answer, was used by V. S. Paskevich [43].

A nonnegative statistic $S(X_1, \cdots, X_n)$ is said to be a tube statistic if it satisfies the following conditions:

(i) $S(x_1, x_2, \cdots, x_n) = 0$ if and only if $x_1 = x_2 = \cdots = x_n$;

(ii) The level surfaces $S(x_1, x_2, \cdots, x_n) = A$ are cylinders with the line $x_1 = x_2 = \cdots = x_n$ as their common axis;

(iii) Any level surface $S(x_1, x_2, \cdots, x_n) = A$ can be obtained from $S(x_1, x_2, \cdots, x_n) = 1$ by a homothetic extension with $x_1 = x_2 \cdots = x_n$ as the axis and the homothetic ratio $f(A)$.

A tube statistic is therefore essentially defined by a geometric property: it provides a partition of the sample space into homothetic cylinders with common axis $X_1 = X_2 = \cdots = X_n$. It is easily seen that a tube statistic is always translation invariant. Many of the important statistics, for instance the sample range or the sample moments of even order, are tube statistics.

We now state the theorem of Paskevich; however, we do not specify certain assumptions which still have to be made concerning the smoothness of the tube statistic. For these we refer the reader to paper [42].

THEOREM A2. *Let $X_1, X_2, \cdots, X_n$ be a sample from a population and assume that the population distribution function $F(x)$ is three times differentiable. A sufficiently smooth tube statistic is independent of the sample mean if, and only if, the population is normal.*

Theorem A2 provides also only a partial answer to the problem of characterizing a normal population by the independence of the sample mean and one sufficiently specialized translation invariant statistic. This is seen if we realize that according to theorem 4.3 the independence of the sample moment of order three and the sample mean characterize the normal population. This cannot be obtained, however, from theorem A2 since the sample moment of order three is a translation-invariant statistic but not a tube statistic.

REFERENCES

[1] D. BASU and R. G. LAHA, "On some characterizations of the normal distribution," *Sankhyā*, Vol. 13 (1954), pp. 359–362.

[2] S. BERNSTEIN, "Sur une propriété caractéristique de la loi de Gauss," *Trans. Leningrad Polytechn. Inst.*, Vol. 3 (1941), pp. 21–22.

[3] R. P. BOAS, "Sur les séries et intégrales de Fourier à coefficients positifs," *C. R. Acad. Sci., Paris*, Vol. 228 (1949), pp. 1837–1838.

[4] H. CRAMÉR, Über eine Eigenschaft der normalen Verteilungsfunktion," *Math. Zeitschrift*, Vol. 41 (1936), pp. 405–414.

[5] ———, *Random Variables and Probability Distributions*, Cambridge, The University Press, 1937.

[6] J. F. DALY, "On the use of the sample range in an analogue of Student's *t*-test," *Annals of Math. Stat.*, Vol. 17 (1946), pp. 71–74.

[7] G. DARMOIS, "Sur une propriété caractéristique de la loi de probabilité de Laplace," *C. R. Acad. Sci., Paris*, Vol. 232 (1951), pp. 1999–2000.

[8] ———, "Analyse générale des liaisons stochastiques," *Revue de l'Inst. Intern. de Stat.*, Vol. 21 (1953), pp. 2–8.

[9] F. FAÀ DI BRUNO, "Note sur une nouvelle formule du calcul différentiel," *Quarterly Jour. of Math.*, Vol. 1 (1857), pp. 359–360.

[10] R. A. FISHER, "Applications of 'Student's' distribution," *Metron*, Vol. 5 (1925), No. 3, pp. 90–104.

[11] M. FRÉCHET, "Généralisations de la loi de probabilité de Laplace," *Ann. de l'Inst. Henri Poincaré*, Vol. 12 (1951), pp. 1–29.

[12] R. C. GEARY, "Distribution of 'Student's' ratio for non-normal samples," *Jour. Roy. Stat. Soc. Suppl.*, Ser. B, Vol. 3 (1936), pp. 178–184.

[13] B. V. GNEDENKO, "On a theorem of S. N. Bernstein," *Izvestiya Akad. Nauk SSSR* (Ser. Mat.), Vol. 12 (1948), pp. 97–100.

[14] B. V. GNEDENKO and A. N. KOLMOGOROV, *Limit Distributions for Sums of Independent Random Variables*, Cambridge, Mass., Addison-Wesley Co., 1954.

[15] M. KAC, "On a characterization of the normal distribution," *Amer. Jour. Math.*, Vol. 61 (1939), pp. 726–728.

[16] T. KAWATA and H. SAKAMOTO, "On the characterization of the normal population by the independence of the sample mean and the sample variance," *Jour. Math. Soc. Japan*, Vol. 1 (1949), pp. 111–115.

[17] R. G. LAHA, "On an extension of Geary's theorem," *Biometrika*, Vol. 40 (1953), pp. 228–229.

[18] ———, "On a characterization of the Gamma distribution," *Annals of Math. Stat.*, Vol. 25 (1954), pp. 784–787.

[19] P. LÉVY, "Propriétés asymptotiques des sommes des variables aléatoires independentes ou enchaînées," *Jour. de Math. Pures et Appliquées*, Vol. 14 (1935), pp. 347–402.

[20] ———, "L'arithmétique des lois de probabilité," *C. R. Acad. Sci., Paris*, Vol. 204 (1937), pp. 80–82.

[21] ———, *Théorie de l'Addition des Variables Aléatoires*, Paris, Gauthier-Villars, 1937.

[22] YU. V. LINNIK, "Linear statistics and the normal distribution law," *Doklady Akad. Nauk SSSR* (N.S.), Vol. 83 (1952), pp. 353–355.

[23] YU. V. LINNIK, "On some identically distributed statistics," *Doklady Akad. Nauk SSSR* (N.S.), Vol. 89 (1953), pp. 9–11.

[24] ———, "Linear forms and statistical criteria. I, II," *Ukrain Mat. Žurnal*, Vol. 5 (1953), pp. 207–243 and 247–290.

[25] E. LUKACS, "A characterization of the normal distribution," *Annals of Math. Stat.*, Vol. 13 (1942), pp. 91–93.

[26] ———, "The stochastic independence of symmetric and homogeneous linear and quadratic statistics," *Annals of Math. Stat.*, Vol. 23 (1952), pp. 442–449.

[27] ———, "Sur une propriété de la loi de Gauss-Laplace," *C. R. Acad. Sci., Paris*, Vol. 238 (1954), pp. 444–445.

[28] ———, "Applications of Faà di Bruno's formula in mathematical statistics," *Amer. Math. Monthly*, Vol. 62 (1955), pp. 340–348.

[29] ———, "Sur une caractérisation de la distribution de Poisson," *C. R. Acad. Sci., Paris*, Vol. 239 (1954), pp. 1114–1116.

[30] E. LUKACS and E. P. KING, "A property of the normal distribution," *Annals of Math. Stat.*, Vol. 25 (1954), pp. 389–394.

[31] E. LUKACS and O. SZÁSZ, "On analytic characteristic functions," *Pacific Jour. Math.*, Vol. 2 (1952), pp. 615–625.

[32] J. MARCINKIEWICZ, "Sur une propriété de la loi de Gauss," *Math. Zeitschrift*, Vol. 44 (1939), pp. 612–618.

[33] E. J. G. PITMAN, "The 'closest estimates' of statistical parameters," *Proc. Camb. Phil. Soc.*, Vol. 33 (1937), pp. 212–222.

[34] G. PÓLYA, "Remarks on characteristic functions," *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1949, pp. 115–123.

[35] D. A. RAIKOV, "On the decomposition of Poisson laws," *Doklady Akad. Nauk SSSR*, Vol. 14 (1937), pp. 9–11.

[36] ———, "On the decomposition of Gauss and Poisson laws," *Izvestiya Akad. Nauk SSSR* (Ser. Mat.), Vol. 2 (1938), pp. 91–124.

[37] ———, "A theorem from the theory of analytic characteristic functions," *Bull. Inst. Mat. Mech. Univ. Tomsk*, Vol. 2, Fasc. II (1938), pp. 8–11.

[38] A. RÉNYI, "On the algebra of distributions," *Publ. Math. Debrecen*, Vol. 1 (1950), pp. 135–149.

[39] V. P. SKITOVICH, "On a property of the normal distribution," *Doklady Akad. Nauk SSSR* (N.S.), Vol. 89 (1953), pp. 217–219.

[40] ———, "Linear forms of independent random variables and the normal distribution," *Izvestiya Akad. Nauk SSSR* (Ser. Mat.), Vol. 18 (1954), pp. 185–200.

[41] A. A. ZINGER, "On independent samples from normal populations," *Uspehi Matem. Nauk* (N.S.), Vol. 6, No. 5 (45), pp. 172–175.

[42] YU. V. LINNIK, "Independent and equally distributed statistics and some statistical and analytical problems connected with them." Mimeographed notes, Calcutta, November 29, 1954.

[43] V. S. PASKEVICH, "On a property of control charts used in running control operations," *Prikladnaya mat. i. mekh.*, Vol. 17 (1953), pp. 49–57.