COLLOQUIUM MATHEMATICUM

VOL. XXXV 1976 FASC. 1

SOME RECENT DIMENSION FREE CHARACTERIZATIONS OF THE NORMAL LAW

BY

A. KUMAR (WAYNE) AND P. K. PATHAK (NEW MEXICO)

DEDICATED TO PROFESSOR J. L. DOOB ON THE OCCASION OF HIS 65TH BIRTHDAY

1. Introduction. Although nearly two decades ago Darmois [3] and Skitovich [24] drew attention to characterizations of the normal law, it was not until recently through the work of Ghurye and Olkin [7], Khatri and Rao [12] and others that many characterizations of normal and other probability distributions have been rendered tractable and established on a sound footing. The literature on characterization problems has now grown substantially. An extensive survey of this field in a recent book by Kagan et al. [11] lists over 180 references. A number of investigators have recently made significant contributions to this subject and many of them have directed their efforts towards a unified dimension-free approach to characterization problems. The object of this article is to present a systematic account of this approach. The motivation for much of our work lies in the pioneering work of Ghurye and Olkin [7], Khatri and Rao [12] and others. It is hoped that our dimension-free approach will shed further insight into the true nature of numerous characterizations of the normal law that have been discovered in recent years. Since this is a review-research paper, a major part of it is being written in a simple expository fashion.

We begin with the solution of certain functional equations which are commonly encountered in characterization problems.

2. Solution of certain functional equations in Hilbert space. Let H, H_1, H_2, etc. denote separable Hilbert spaces. Let φ be a given complex-valued function on H. If h is an element of H, we define the *difference operator with increment* h, $\Delta_h\varphi\colon H\to C$, by the equation

$$(\Delta_h\varphi)(x) = \varphi(x+h)-\varphi(x). \tag{2.1}$$

In an analogous manner the *n-th difference operator* $\Delta_{h(1),\dots,h(n)}$ is defined inductively as follows:

$$(\Delta_{h(1),\dots,h(n)}\varphi)(x) = \big(\Delta_{h(n)}(\Delta_{h(1),\dots,h(n-1)}\varphi)\big)(x). \tag{2.2}$$

In terms of the difference operator Δ we have ([16] and [17])

Definition 2.1. A continuous function P: $H \to C$ is a *polynomial of degree less than or equal to n* if every $(n+1)$-st order difference of P vanishes; if, in addition, for some $h(1), \dots, h(n)$, the n-th order difference $\Delta_{h(1),\dots,h(n)}P \neq 0$, P is said to be a *polynomial of degree n*.

By convention, a non-zero constant function is a *polynomial of degree zero* while the zero function $P \equiv 0$ is a polynomial of degree -1.

The following lemmas play a central role in characterization problems:

LEMMA 2.1. *Suppose that* P: $H_1 \times H_2 \to C$ *is continuous and there exists an integer* n *such that for each* y *the function* $x \to P(x, y)$ *is a polynomial of degree less than or equal to* m. *Furthermore, suppose that for each element* v *of* H_2 *there is an element* $w(v)$ *such that the function* $Q(x, y) = P(x+w, y+v) - P(x, y)$ *is a polynomial on* $H_1 \times H_2$ *and the degree of* Q *is less than or equal to* n *for all* v *in* H_2. *Then* P *is a polynomial on* $H_1 \times H_2$ *of degree less than or equal to* $\max(m, n+1)$.

An elementary proof of this lemma is given in [8]. Actually the lemma furnishes the best possible degree for P; for an example where the degree of P is attained, let $P(x, y) = P_m(x) + R_{n+1}(y)$ with P_m and R_{n+1} being polynomials of degree m and $n+1$, respectively.

LEMMA 2.2. *Let* P: $H_1 \times H_2 \to C$ *and* Q: $H_1 \times H_2 \to C$ *be continuous and suppose that, for each* y, *the functions* $x \to P(x, y)$ *and* $x \to Q(x, y)$ *are polynomials of degree less than or equal to* p *and* q, *respectively. Let* $E(x, y)$ *be a polynomial of degree* m *on* $H_1 \times H_2$. *Let there be given linear operators* A_j: $H_1 \to H$ *and* B_j: $H_2 \to H$ *with* $\operatorname{range}(A_j) \subset \operatorname{range}(B_j)$ *and continuous functions* φ_j: $H \to C$, $1 \leqslant j \leqslant n$, *such that, for all* $(x, y) \in H_1 \times H_2$,

$$\sum_1^n \varphi_j(A_j x + B_j y) = P(y, x) + Q(x, y) + E(x, y). \tag{2.3}$$

Then P *is a polynomial on* $H_1 \times H_2$ *of degree less than or equal to* $\max(m, n+p+q)$.

An elementary proof of the lemma is given in [8]. An example in which the stated degree of P is attained can be easily constructed by letting $n = 1$, $\varphi(x) = x$, $A = B = I$ and $Q(x, y) = C_p(x)D_q(y)$, where C and D are polynomials of degree p and q, respectively.

COROLLARY 1. *Suppose that in* (2.3) $\operatorname{range}(A_j) = \operatorname{range}(B_j)$. *Then* $\sum_1^n \varphi_j(A_j x + B_j y)$ *is a polynomial on* $H_1 \times H_2$ *of degree less than or equal to* $\max(m, p+q+n)$.

LEMMA 2.3. *Let φ_j: $H \to C$, $j = 1, \dots, n$, be continuous functions and suppose that there exist linear operators A_j: $H \to H$ such that the function $\sum_1^n \varphi_j(A_j x)$ is a polynomial of degree less than or equal to m on H. Define the mapping (A_i, A_j): $x \to (A_i x, A_j x)$ which maps H into $H \times H$. Suppose r is such that, for each (i, j) satisfying $1 \leqslant i \leqslant r < j \leqslant n$, the range of $(A_i, 0)$ is contained in the range of (A_i, A_j). Then $\sum_1^r \varphi_j(A_j x)$ is a polynomial of degree less than or equal to* $\max(m, n-2)$.

Proof. Although this lemma is correctly stated in [8], its proof contains an error. It is, therefore, worthwhile to furnish a correct proof of the lemma here. Suppose that

$$\sum_1^n \varphi_j(A_j x) = P(x),$$

where $P(x)$ is a polynomial of degree less than or equal to m. If $r = n$, the result holds trivially. Let $r < n$. Let $z_j \in \ker(A_j)$, $r < j \leqslant n$, be the vectors in H chosen arbitrarily. Then

$$\sum_1^r \Delta_{z_{r+1}} \dots \Delta_{z_n} \varphi_i(A_i x) = P(x, z_{r+1}, \dots, z_n), \tag{2.4}$$

where $P(x, z_{r+1}, \dots, z_n)$ is a polynomial of degree less than or equal to m in $(x, z_{r+1}, \dots, z_n)$. Clearly, (2.4) can be written as

$$\sum_1^r (\Delta_{z_{r+2}} \dots \Delta_{z_n} \varphi_i)(A_i x + A_i z_{r+1}) = Q(x) + P(x, z_{r+1}, \dots, z_n), \tag{2.5}$$

$$\text{where } Q(x) = \sum_1^r \Delta_{z_{r+2}} \dots \Delta_{z_n} \varphi_i(A_i x).$$

Since $\text{range}(A_i, 0) \subset \text{range}(A_i, A_j)$, $1 \leqslant i \leqslant r < j \leqslant n$, it follows that

$$\{A_i x\colon x \in H\} = \{A_i z_j\colon z_j \in \ker(A_j)\} \quad \text{for } 1 \leqslant i \leqslant r < j \leqslant n.$$

Also, given $z_{r+2}, \dots, z_n$, $P(x, z_{r+1}, z_{r+2}, \dots, z_n)$ is a polynomial in (x, z_{r+1}) of degree less than or equal to $m-(n-r)+1$. So from Lemma 2.2, for given $(z_{r+2}, \dots, z_n)$, $Q(x)$ is a polynomial in x of degree less than or equal to $\max(m-(n-r)+1, r-1) = k$, say. Therefore we can write

$$\sum_1^r (\Delta_{z_{r+2}} \dots \Delta_{z_n} \varphi_i)(A_i x) = P_k(x, z_{r+2}, \dots, z_n). \tag{2.6}$$

We now show that (2.6) implies that $P_k(x, z_{r+2}, z_{r+3}, \dots, z_n)$ for given $(z_{r+3}, \dots, z_n)$ is a polynomial on $H \times \ker(A_{r+2})$ of degree less than or equal to $k+1$. Let h be an arbitrary element of $\ker(A_{r+2})$. Then it is easily seen that

$$(2.7)\qquad P_k(x-h, z_{r+2}+h, z_{r+3}, \dots, z_n) - P_k(x, z_{r+2}, \dots, z_n) = P_k(x-h, h, z_{r+3}, \dots, z_n),$$

where, for given $z_{r+3}, \dots, z_n$ and h, the right-hand side is a polynomial in x and z_{r+2} (trivially!) of degree less than or equal to k. So it follows from Lemma 2.1 that, for given $z_{r+3}, \dots, z_n$,

$$(2.8)\qquad P_k = \sum_1^r (\Delta_{z_{r+2}} \cdots \Delta_{z_n} \varphi_i)(A_i x)$$

is a polynomial in (x, z_{r+2}) of degree less than or equal to

$$k+1 = \max\big(m-(n-2)+2, r\big).$$

Successive iteration of process (2.6) through (2.8) then yields the desired result.

Khatri and Rao [12] established this lemma for finite-dimensional vector spaces with the A_i as matrices which satisfy

$$\operatorname{rank}\begin{pmatrix} A_i \\ A_j \end{pmatrix} = \operatorname{rank}(A_i) + \operatorname{rank}(A_j), \qquad 1 \leqslant i \leqslant r < j \leqslant n.$$

It is easy to see that this condition on matrices implies that if $A_i x = b$ for some x, then the system of equations

$$\begin{pmatrix} A_i \\ A_j \end{pmatrix} y = \begin{pmatrix} b \\ 0 \end{pmatrix}$$

admits a solution so that $\operatorname{range}(A_i, 0) \subset \operatorname{range}(A_i, A_j)$ showing that their condition is the appropriate analogue of the condition on the operators A_i in Lemma 2.3.

COROLLARY 1. *Suppose that the operators A_i of Lemma* 2.3 *satisfy* $\operatorname{range}(A_i, 0) \subset \operatorname{range}(A_i, A_j)$ *for $i \neq j$. Then each $\varphi_i(A_i x)$ is a polynomial of degree less than or equal to* $\max(m, n-2)$.

COROLLARY 2. *Consider the functions and operators of Lemma* 2.2 *and suppose that they satisfy* (2.3). *Moreover, suppose that* $\operatorname{range}(A_i) = \operatorname{range}(B_i)$, $1 \leqslant i \leqslant n$, *and there is an r such that, for $1 \leqslant i \leqslant r < j \leqslant n$, the range of $(A_i x + B_i y, 0)$ is contained in the range of $(A_i x + B_i y, A_j x + B_j y)$. Then $\sum_1^r \varphi_i(A_i x + B_i y)$ is a polynomial on $H \times H$ of degree less than or equal to* $\max(m, n+p+q)$.

COROLLARY 3. *Let B_j, $1 \leqslant j \leqslant n$, be surjective operators and suppose that $(B_i - B_j)$ for each $i \neq j$ is also surjective. Let φ_j be continuous functions, $1 \leqslant j \leqslant n$, which satisfy*

$$\sum_1^n \varphi_i(x + B_i y) = P(x) + Q(y). \tag{2.9}$$

Then each φ_i is a polynomial of degree less than or equal to n.

This corollary is an immediate consequence of the preceding corollary. The finite-dimensional version of Corollary 3 is due to Ghurye and Olkin [7].

COROLLARY 4. *Let φ_j, $1 \leqslant j \leqslant n$, be n given continuous functions in H. Let A_j, B_j, C_j, $1 \leqslant j \leqslant n$, be linear operators in H. Let $P(x, y, z)$, $Q(x, y, z)$, $R(x, y, z)$ and $E(x, y, z)$ be given functions in $H \times H \times H$. Suppose that P, Q and R for each z are polynomials in $H \times H$ of degree p, q and r, respectively. Suppose that E is a polynomial of degree less than or equal to m in $H \times H \times H$. Consider now the functional equation*

$$\begin{aligned} \sum_1^n \varphi_j(A_j x + B_j y + C_j z) \\ = P(y, z, x) + Q(z, x, y) + R(x, y, z) + E(x, y, z). \end{aligned} \tag{2.10}$$

Then the following assertions hold:

(a) *If* $\mathrm{range}(A_j x) \subset \mathrm{range}(B_j y + C_j z)$, $1 \leqslant j \leqslant n$, *then P is a polynomial on the space $H \times H \times H$ of degree less than or equal to* $\max\{m, n+p+q, n+p+r\}$.

(b) *If, for each j, $1 \leqslant j \leqslant n$,*

$$\mathrm{range}(A_j x) \subset \mathrm{range}(B_j y + C_j z),$$

$$\mathrm{range}(B_j y) \subset \mathrm{range}(A_j x + C_j z) \quad \textit{and} \quad \mathrm{range}(C_j z) \subset \mathrm{range}(A_j x + B_j y),$$

then

$$\sum_1^n \varphi_j(A_j x + B_j y + C_j z)$$

is a polynomial in the space $H \times H \times H$ of degree less than or equal to $\max\{m, n+p+q, n+p+r\}$.

(c) *Suppose that, in addition to the hypotheses of* (b), *there is an r such that, for $1 \leqslant i \leqslant r < j \leqslant n$,*

$$\mathrm{range}\{A_i x + B_i y + C_i z, 0\} \subset \mathrm{range}\{A_i x + B_i y + C_i z, A_j x + B_j y + C_j z\}.$$

Then

$$\sum_1^r \varphi_i(A_i x + B_i y + C_i z)$$

is a polynomial on the space $H \times H \times H$ *of degree less than or equal ot* $\max\{m, n+p+q, n+p+r\}$.

(d) *Suppose that, in addition to the hypotheses of* (b), *for each* $i \neq j$,

$$\text{range}\{A_i x + B_i y + C_i z, 0\} \subset \text{range}\{A_i x + B_i y + C_i z, A_j x + B_j y + C_j z\}.$$

Then each $\varphi_i(A_i x + B_i y + C_i z)$ *is a polynomial in* $H \times H \times H$ *of degree less than or equal to* $\max\{m, n+p+q, n+p+r\}$.

This corollary has an obvious extension to $H_1 \times \ldots \times H_k$ for any integer k. It is perhaps worthwhile noting that the corollary is the natural extension of a similar result of Khatri and Rao [12] in finite-dimensional vector spaces.

Lemma 2.3 and its corollaries furnish best possible results of their kind under their respective hypotheses; examples to this effect can be easily constructed. In applications there are, however, cases in which one has much more information about the underlying functions and operators than the preceding lemmas and corollaries make use of. The added information can of course be used to improve the degrees of polynomials in preceding results. The most effective technique for doing so is to carry out an iterative procedure of differencing with a minimum number of steps which eliminates from the functional equation all the unnecessary functions φ_j with the exception of those whose degrees are desired. For illustration consider the following example:

Example. Let φ_j, $1 \leqslant j \leqslant 6$, be six continuous functions and suppose that they satisfy the equation

$$\varphi_1(x) + \varphi_2(y) + \varphi_3(z) + \varphi_4(x+y) + \varphi_5(x+z) + \varphi_6(y+z) = 0. \tag{2.11}$$

An application of Corollary 1 to Lemma 2.3 shows that each φ_i is a polynomial of degree less than or equal to 4. Now to obtain a more precise estimate of φ_1, say, replace x by $x+h$, y by $y-h$ and z by $z-h$ in (2.11) and obtain the following equation:

$$\varphi_1(x+h) + \varphi_2(y-h) + \varphi_3(z-h) + \varphi_4(x+y) + \varphi_5(x+z) + \\ + \varphi_6(y+z-2h) = 0. \tag{2.12}$$

On subtracting (2.12) from (2.11) we find that $\Delta_h \varphi_1(x)$ is free of x so that φ_1 is a linear function.

In an analogous manner, an equation similar to (2.11) in p variables $x_1, \ldots, x_p$ and $p(p+1)/2$ unknown functions $\varphi_1, \ldots, \varphi_{p(p+1)/2}$ can be shown to imply that each of the φ_i must be linear. This last observation has been used by Rao ([11] and [23]) to show that the joint probability distribution of p suitably chosen linear functions of $p(p+1)/2$ independent random variables determines the probability distribution of each variable up to a change of location.

The "technique of elimination of functions", as illustrated above, has been used in the literature in many different ways to solve functional equations commonly encountered in applications. It would be impossible to even attempt to present all such applications here. Nonetheless we furnish below two lemmas which it is hoped would demonstrate the strength of this technique.

LEMMA 2.4. *Let φ_j: $H \to C$, $j = 1, \ldots, r+2s$, be continuous functions and suppose that there exist linear operators A_j: $H \to H$ such that the function $\sum_{1}^{r+2s} \varphi_j(A_j x)$ is a polynomial of degree less than or equal to m. Define the mapping (A_i, A_j, A_k): $x \to (A_i x, A_j x, A_k x)$ which maps H into $H \times H \times H$. Suppose that* $\operatorname{range}(A_i, 0, 0) \subset \operatorname{range}(A_i, A_j, A_{j+s})$ *for* $1 \leqslant i \leqslant r$ *and* $r+1 \leqslant j \leqslant r+s$. *Then $\sum_{1}^{r} \varphi_i(A_i x)$ is a polynomial of degree less than or equal to* $\max(m, r+s-2)$.

Proof. Let $z_j \subset \ker(A_j) \cap \ker(A_{j+s})$, $r+1 \leqslant j \leqslant r+s$, be vectors in H chosen arbitrarily. Then, by the hypotheses of the lemma, it follows that

$$(2.13) \qquad \sum_{1}^{r} \Delta_{z_{r+1}} \ldots \Delta_{z_{r+s}} \varphi_i(A_i x) = P(x, z_{r+1}, \ldots, z_{r+s}),$$

where P is a polynomial in $(x, z_{r+1}, \ldots, z_{r+s})$ on

$$H \times [\ker(A_{r+1}) \cap \ker(A_{r+s})] \times \ldots \times [\ker(A_{r+s}) \cap \ker(A_{r+2s})]$$

of degree less than or equal to m. An approach similar to that of Lemma 2.3 shows that $\sum_{1}^{r} \varphi_i(A_i x)$ is a polynomial of degree less than or equal to $\max(m, r+s-2)$.

LEMMA 2.5. *Let φ_j: $H \to C$, $j = 1, \ldots, n$, be continuous functions and suppose that the functions satisfy the equation*

$$\sum_{1}^{n} \varphi_j(A_j x) = 0,$$

where the A_j are linear operators. Further suppose that there is an r, $2 \leqslant r \leqslant n$, such that

$$\operatorname{range}(A_1, 0, \ldots, 0) \subset \operatorname{range}(A_1, A_2, \ldots, A_r)$$

and

$$A_1 A'_{r+1} = \ldots = A_1 A'_n = 0,$$

where A' denotes the adjoint of A. Then $\varphi_1(A_1 x)$ is linear.

Proof. Let $K = \ker(A_2) \cap \ldots \cap \ker(A_r)$. The equation

$$\sum_{1}^{n} \varphi_j(A_j x) = 0$$

then yields

$$\Delta_z\varphi_1(A_1x) = -\sum_{r+1}^{n}\Delta_z\varphi_j(A_jx) \quad \text{for each } z\in K. \tag{2.14}$$

Now let P denote the orthogonal projection on $[\ker(A_1)]^{\perp}$. Then $\ker(P) = \ker(A_1)$ and $A_1P = A_1$. It also follows from the hypothesis of the lemma that $A_jP = 0$ for $j = r+1, \ldots, n$. We thus have from (2.14)

$$\Delta_z\varphi_1(A_1x) = \Delta_z\varphi_1(A_1Px) = -\sum_{r+1}^{n}\Delta_z\varphi_j(A_jPx) = -\sum_{r+1}^{n}\Delta_z\varphi_j(0) \quad \text{for each } z\in K, \tag{2.15}$$

so that $\Delta_z\varphi_1(A_1x)$ is constant in x. Since $\{A_1x\colon x\in H\} = \{A_1z\colon z\in K\}$ in our case, it follows that $\varphi_1(A_1x)$ is linear in x.

Before concluding this section a few remarks are perhaps in order. First, it can be easily seen that Lemmas 2.4 and 2.5 have a number of corollaries which are analogous in nature and proof to those of Lemma 2.3. For reasons of brevity, we leave their verification to the reader. The second remark is concerning the extension of these results to other linear spaces, e.g. Banach space, locally compact groups and so forth. The results presented in this section carry over word for word to these spaces except Lemma 2.5, a suitable extension of which is also possible. Finally, in many applications one finds that functional equations considered in preceding lemmas are only satisfied in a neighborhood of the origin and in such cases our results remain valid but only in a certain neighborhood of the origin.

We turn now to certain characterizations of the normal law.

3. Characterizations of the normal law. In what follows we first furnish a few well-known results concerning characterizations of the normal law on the line. Not only do these results motivate characterizations of the normal law in linear spaces but are also useful in establishing many such multivariate generalizations.

Definition 3.1. A real-valued function $U(s,t)$ of the complex-variable $z = s+it$ is called *harmonic* if U_{ss} and U_{tt} are continuous and $U_{ss}+U_{tt} = 0$.

LEMMA 3.1. *A real-valued function $U(s,t)$ is harmonic iff U is the real part of an analytic function.*

LEMMA 3.2. *Let $f(z)$ be a non-vanishing analytic function. Then* $\log|f(z)|$ *is harmonic.*

LEMMA 3.3. *Let $U(z)$ be a harmonic function and suppose that $U(z) = O(|z|^2)$. Then $U(z)$ is a polynomial in z of degree less than or equal 2.*

These lemmas are standard results in harmonic functions and can be looked up, for example, in Ahlfors [1].

LEMMA 3.4. *Let X and Y be two independent random variables and suppose that the moment generating function $M_S(t)$ of $S = X+Y$ exists for $|t| < T$. Then the moment generating function $M_X(t)$ of X exists for $|t| < T$ and satisfies, for some $k > 0$ and $a > 0$,*

$$\frac{1}{k} e^{-a|t|} \leqslant M_X(t) \leqslant k e^{a|t|} M_S(t) \tag{3.1}$$

(cf. Ramachandran [21] for a proof).

LEMMA 3.5. *Let φ be a non-vanishing characteristic function (c.f.) on the real line. Then there exists a unique continuous function λ with $\lambda(0) = 0$ such that $\varphi(t) = \exp[\lambda(t)]$ for each t.*

See Chung [2] for a proof. An analogous method of proof yields the following extension in a Hilbert space [9]:

LEMMA 3.6. *Let f be a complex-valued non-vanishing uniformly continuous function defined on a real Hilbert space H so that $f(0) = 1$. Then there is a unique continuous function λ on H such that $\lambda(0) = 0$ and $f(y) = \exp[\lambda(y)]$, $y \in H$.*

LEMMA 3.7 (Cramér). *Let X and Y be two independent real-valued random variables and suppose that $U = X+Y$ has the normal distribution. Then X has the normal distribution.*

Proof. Since $U = X+Y$ has the normal distribution, its moment generating function admits the form $M_U(s) = \exp[as-bs^2]$, where $b > 0$. By Lemma 3.4 the moment generating function of X satisfies the inequality

$$M_X(s) \leqslant k \exp[a|s|-bs^2], \tag{3.2}$$

and a similar result holds for $M_Y(s)$. For $z = s+it$, define $M_X(z) = \mathrm{E}[e^{zX}]$, and $M_Y(z)$ and $M_U(z)$ similarly. Then, by Lemma 3.4,

$$|M_X(z)| \leqslant M_X(s) \leqslant k \exp[a|s|-bs^2]$$

and, similarly,

$$|M_Y(z)| \leqslant k \exp[a|s|-bs^2].$$

Since $M_X(z) = M_U(z)/M_Y(z)$, it follows that

$$\frac{1}{k} \exp[-a|s|] \leqslant |M_X(z)| \leqslant k \exp[a|s|-bs^2], \quad \text{where } s = \mathrm{Re}(z). \tag{3.3}$$

Since the moment generating function of X exists, it follows that $M_X(z)$ is analytic; it is also non-vanishing. Thus $\log|M_X(z)|$ is harmonic and, by virtue of (3.3), $\log|M_X(z)| = O(|z|^2)$. So $\log|M_X(z)| = P(z)$, where

$P(z)$ is a quadratic polynomial in z. Hence the moment generating function $M_X(s) = \exp[P(s)]$. So X is normal.

The reader may perhaps wonder the need for this proof of such an ancient theorem in the theory of characterizations. The main object of our new proof is to demonstrate the use of harmonic functions in characterization problems. We hope that our proof might motivate applications of harmonic functions to other characterization problems.

The theorem of Cramér (Lemma 3.7) on decomposition of the normal distribution has been extended in several directions, e.g. decompositions theorems are available for the binomial, the Poisson and normal distributions and so forth (cf. Ramachandran [21]). Similar decomposition results are not fully known for stable distributions, although it is known that if the sum of two independent random variables has a stable distribution, then each variable is not necessarily stable. A partial result in this direction is the following

THEOREM 3.1. *Let X_1, X_2, X_3 be independent random variables and suppose that $sX_1+(s+t)X_2+tX_3$ for each (s, t) has a stable distribution with exponent a. Then each X_i $(i = 1, 2, 3)$ is stable with the same exponent a.*

Proof. We refer the reader to [11] for the definition of stable distribution with exponent a. Let

$$\varphi_i(\cdot) = \log f_i(\cdot) - \log g_i(\cdot),$$

where $\log f_i(\cdot)$ and $\log g_i(\cdot)$ denote two alternative log characteristic functions of X_i $(i = 1, 2, 3)$. (Since $sX_1+(s+t)X_2+tX_3$ has a non-vanishing c.f., each X_i has a non-vanishing c.f., so that $\log f_i$ and $\log g_i$ are unambiguously defined.) A little consideration now shows that

$$\varphi_2(s+t) = -[\varphi_1(t)+\varphi_2(s)]. \tag{3.4}$$

By Lemma 2.2 it follows that each φ_i is linear. Consequently, the distribution of each random variable is determined up to a change of location. Since the sum $sX_1+(s+t)X_2+tX_3$ is always stable with exponent a when the X_i are stable with exponent a, it follows that the X_i must have stable distribution with exponent a.

In a similar fashion we can establish the following extension. Let $X_1, \ldots, X_{n(n+1)/2}$ be independent random variables and suppose that, for each $(t_1, \ldots, t_n)$, the distribution of

$$t_1X_1+t_2X_2+\ldots+t_nX_n+(t_1+t_2)X_{n+1}+\ldots+(t_{n-1}+t_n)X_{n(n+1)/2}$$

is stable with exponent a. Then each X_i has a stable distribution with the same exponent.

LEMMA 3.8 (Marcinkiewicz). *Suppose that the c.f. φ of a real-valued random variable X admits the form $\varphi(t) = \exp[P(t)]$, where $P(t)$ is a polynomial in t in a neighborhood of $t = 0$. Then X is normally distributed* (cf. [21] and Lemma 2.4.3 in [11] for a proof).

A simple extension of Lemma 3.8 is the following corollary:

COROLLARY 1. *Let X be a random variable and suppose that the function $\varphi(z) = \mathrm{E}[\exp(zX)]$ exists for all $z = s+it$ and is analytic and non-vanishing. Also suppose that $\log|\varphi(z)|$ has polynomial growth in z. Then X is normally distributed.*

With these preliminaries we turn now to dimension-free characterizations of the normal law. In order to furnish a unified approach to characterization problems we assume that, unless stated to the contrary, the random variables under study take values in a given real separable Hilbert space H. We keep in the back of our minds a system of coordinates in H, i.e. a complete orthonormal sequence $\{e_n\colon n \geqslant 1\}$ of vectors in H. For each $x \in H$, we denote by x_n the n-th coordinate of x, i.e. $x_n = (x, e_n)$. In our set up we identify a real-valued random variable X with Xe_1 in H; similarly, a set of finitely many random variables $(X_1, \ldots, X_k)$ will be identified with the random variable $Y_1e_1+\ldots+Y_ke_k$ in H and so forth. Finally, an H-valued random variable X is said to have the *normal distribution* if, for each $y \in H$, the real-valued random variable (y, X) has the normal distribution on the line. We now begin our characterizations with certain extensions of the Darmois-Skitovich theorem

THEOREM 3.2. *Let $X_1, \ldots, X_n$ be independent random variables. Let A_{ij}, $1 \leqslant j \leqslant n$, be bounded linear operators in H. Suppose that the m random variables*

$$Y_i = \sum_{j=1}^{m} A_{ij}' X_j, \qquad 1 \leqslant i \leqslant m,$$

are mutually independent and, for each j, $1 \leqslant j \leqslant n$,

$$\{A_{1j}t_1\colon t_1 \in H\} \subset \{A_{2j}t_2+\ldots+A_{mj}t_m\colon t_i \in H,\ i = 2, \ldots, m\}.$$

Then Y_1 has a normal distribution.

Proof. Let $\varphi_j(t) = \mathrm{E}\big[\exp[i(t, X_j)]\big]$ denote the c.f. of X_j. Since φ_j does not vanish in a neighborhood of $t = 0$, we can take its logarithm ψ_j in a certain neighborhood of $t = 0$. Now the mutual independence of the Y_i, $1 \leqslant i \leqslant m$, yields in a neighborhood of $(t_1, \ldots, t_m) = 0$ the equation

$$\sum_{j=1}^{n} \psi_j(A_{1j}t_1+\ldots+A_{mj}t_m) = C_1(t_1)+\ldots+C_m(t_m), \tag{3.5}$$

where $C_i(t)$ denotes the logarithm of the c.f. (l.c.f.) of Y_i, $1 \leqslant i \leqslant m$. By assumption

$$\text{range}(A_{1j}t_1) \subset \text{range}(A_{2j}t_2+\ldots+A_{mj}t_m).$$

Therefore, by Corollary 4 to Lemma 2.3, it follows that $C_1(t_1)$ is a polynomial in a neighborhood of $t_1 = 0$. For fixed $x \in H$, a direct calculation shows that the c.f. of (x, Y_1) at t, $-\infty < t < \infty$, is given by $\exp[C(tx)]$ for small $|t|$. By Marcinkiewicz's theorem we then see that, for each x, the real random variable (x, Y_1) is normal. So Y_1 is normal.

COROLLARY 1. *Suppose that, for each i, $1 \leqslant i \leqslant m$,*

$$\text{range}(A_{ij}t_i) \subset \text{range}(A_{1j}t_1+A_{2j}t_2+\ldots+A_{mj}t_m-A_{ij}t_i), \qquad 1 \leqslant j \leqslant n.$$

Then each Y_i is normally distributed.

Theorem 3.2 and this Corollary 1 are dimension-free extensions of similar results due to Ghurye and Olkin [7], Khatri and Rao [12] and others. Results of this kind have been used in the literature to establish normality of the basic variables X_j, $1 \leqslant j \leqslant n$, *via* Cramér's theorem. For example, if Y_1 is normal, Cramér's theorem implies that then so is $A_{11}'X_1$. In that case $(A_{11}t, X_1)$ is normal for each t, and if A_{11} happens to be surjective, then X_1 has a normal distribution. (Actually, to assert normality of X_1 it suffices if the range of A_{11} contains the range of X_1). Consider the other extreme case in which the operators satisfy the hypotheses of Corollary 1. Then all the Y_i, $1 \leqslant i \leqslant m$, are normal. The independence of Y_i implies that, for each $t_1, \ldots, t_m$, the real random variable $(t_1, Y_1)+\ldots+(t_m, Y_m)$ has a normal distribution. By Cramér's theorem it then follows that $(A_{11}t_1+\ldots+A_{1m}t_m, X_1)$ is normal for each $t_1, \ldots, t_m$. So if $A_{11}t_1+\ldots+A_{1m}t_m$ is surjective, then X_1 has a normal distribution. When applied to finite-dimensional random variables Theorem 3.2 yields the following result due to Khatri and Rao [12]:

COROLLARY 2. *Let X_j be a p_j-vector variable, $1 \leqslant j \leqslant n$, and the operator A_{ij}' a $(q_i \times p_j)$-matrix. Let $A_j = (A_{1j}, \ldots, A_{mj})$, and let $A_j(k)$ be the matrix obtained by deleting the k-th partition from A_j. Suppose that* $\text{rank}(A_j(k)) = p_j$ *for all k and j. Then X_j has a p_j-variate normal distribution.*

For a proof, note that under the hypotheses of the corollary, for each i,

$$\text{range}(A_{ij}t_i) \subset \text{range}(A_{1j}t_1+\ldots+A_{mj}t_m-A_{ij}t_i) \quad \text{for all } i \text{ and } j.$$

Also, for each j, $A_{1j}t_1+\ldots+A_{mj}t_m$ is surjective. Conclusions of Corollary 2 are now easily seen to be consequences of Corollary 1.

In Chapter 14 of their book, Kagan et al. (cf. [11], p. 460) mention as unsolved problems extensions of Corollary 2 to Hilbert space-valued and more general random variables and also to a denumerable number of vectors. We consider our Theorem 3.2 and Corollary 1 as the appropriate extension of their result. We mention without going into details that Theorem 3.2 is valid, for example, in any Banach space and also has an appropriate extension to a denumerable number of random variables.

We conclude this section on a slightly different note by partially solving a conjecture communicated to us by Professor H. Kesten.

CONJECTURE. Let X, Y, and Z be random variables such that X is independent of the pair (Y, Z). Suppose that $(X+Y)$ is independent of $(X+Z)$. Then X is normally distributed.

Let $\psi_1(t)$ denote the l.c.f. of X and $\psi_2(t_1, t_2)$ the l.c.f. of (Y, Z) in a neighborhood of the origin. Independence of $(X+Y)$ and $(X+Z)$ now implies that

$$\psi_1(t+t_2)-\psi_1(t_1)-\psi_1(t_2) = \psi(t_1, 0)+\psi(0, t_2)-\psi_2(t_1, t_2)$$

in a neighborhood of $t_1 = 0$, $t_2 = 0$.

To assert normality of X, we must show that the above equation implies that ψ_1 is a polynomial. At the present time we cannot achieve this without added restrictions on $\psi_2(t_1, t_2)$. For example, normality of X_1 is obvious if $\psi(t_1, t_2)$ can be expressed as the sum of functions each of which depends on (t_1, t_2) only through a linear function in t_1 and t_2. We transmit the conjecture here with the hope that interested readers may be able to prove it without any additional restrictions.

4. Characterization of vectors with linear structure. We now turn to a dimension-free discussion of Rao's results [22] on random vectors with linear structure. We first present a few preliminary results and definitions in this connection.

Definition 4.1. A Hilbert space-valued random variable X is said to admit a *linear structure* if there exist non-degenerate independent real-valued random variables $U_1, \ldots, U_k$ and scalars $\beta_0, \beta_1, \ldots, \beta_k$ in H such that

$$X = \beta_0+U_1\beta_1+\ldots+U_k\beta_k, \tag{4.1}$$

where it is assumed, without loss of generality, that no two β_i's are multiples of each other.

Any two representations of X, say, (4.1) and $X = \gamma_0+V_1\gamma_1+\ldots+V_l\gamma_l$ are *equivalent* if each β_i is a multiple of γ_j, and each γ_j is a multiple of some β_i. Finally, the random variable X is said to have a *unique structure* if all linear structures of X are equivalent.

An immediate consequence of Definition 4.1 is that if X admits two representations, say

$$X = \beta_0+U_1\beta_1+\ldots+U_k\beta_k = \gamma_0+V_1\gamma_1+\ldots+V_l\gamma_l, \tag{4.2}$$

then $\langle\beta_1, \ldots, \beta_k\rangle = \langle\gamma_1, \ldots, \gamma_l\rangle$, where $\langle\beta_1, \ldots, \beta_k\rangle$ denotes the subspace spanned by the β_i. This result also implies that

$$(\beta_0 - \gamma_0) \epsilon \langle\beta_1, \ldots, \beta_k\rangle = \langle\gamma_1, \ldots, \gamma_l\rangle,$$

so that β_0 and γ_0 can be eliminated from (4.2) by suitable subtraction.

The above definitions and results are due to Rao [22] except that they have now been paraphrased in Hilbert space terminology.

The following lemma plays a central role in the study of random variables with linear structure:

LEMMA 4.1. *Consider a random variable X with the following two linear structures:*

$$X = U_1\beta_1 + \ldots + U_k\beta_k = V_1\gamma_1 + \ldots + V_l\gamma_l. \tag{4.3}$$

Suppose that no γ_i is proportional to β_1. Then U_1 has a normal distribution on the line.

Proof. Let φ_i, $1 \leqslant i \leqslant k$, and ψ_j, $1 \leqslant j \leqslant l$, denote the l.c.f. of U_i and of V_j, respectively, in a neighborhood of the origin. Then considering the logarithm of characteristic functional of X yields the functional equation

$$\varphi_1\big((\beta_1, t)\big) + \ldots + \varphi_k\big((\beta_k, t)\big) - \psi_1\big((\gamma_1, t)\big) - \ldots - \psi_l\big((\gamma_l, t)\big) = 0 \tag{4.4}$$

in a neighborhood of $t = 0$. It is clear from the hypotheses of the lemma that the range of $\big((\beta_1, \cdot), 0\big)$: $t \to \big((\beta_1, t), 0\big)$ is contained in the range of (β_1, β_i): $t \to \big((\beta_1, t), (\beta_i, t)\big)$, and in the range of (β_1, γ_j): $t \to \big((\beta_1, t), (\gamma_j, t)\big)$. So, by Lemma 2.3, $\varphi\big((\beta_1, t)\big)$ is a polynomial in a neighborhood of $t = 0$. An application of Lemma 3.8 now shows that U_1 has a normal distribution on the line.

COROLLARY 1. *Suppose that no β_i, $1 \leqslant i \leqslant k$, is proportional to any γ_j, $1 \leqslant j \leqslant l$. Then X has a normal distribution.*

COROLLARY 2. *Suppose that β_1 is proportional to γ_1. Then the l.c.f. of U_1 and of V_1 differ by a polynomial of degree less than or equal to $k + l - 3$ in a neighborhood of the origin.* (Clear by Lemma 2.3.)

An immediate consequence of the above results is that if X admits two representations given by (4.3), then the following dichotomy holds for each i, $1 \leqslant i \leqslant m$. Either β_i is not proportional to any γ_j (in that case U_i has a normal distribution on the line) or β_i is proportional to some γ_i (in that case the l.c.f. of U_i and of V_j differ by a polynomial in a neighborhood of $t = 0$).

COROLLARY 3. *Let r, $1 \leqslant r \leqslant k$, be such that $U_1, \ldots, U_r$ are non-normal variables while $U_{r+1}, \ldots, U_k$ are normal. Then every structure of X admits the representation*

$$X = V_1\beta_1 + \ldots + V_r\beta_r + V_{r+1}\gamma_{r+1} + \ldots + V_l\gamma_l, \tag{4.5}$$

where the V_i, $1 \leqslant i \leqslant r$, are non-normal variables and the remaining V_j, $r+1 \leqslant j \leqslant 1$, are normal variables.

Furthermore, in a neighborhood of $t = 0$ the equation

$$\sum_1^r f_i\big((\beta_i, t)\big) = P_2(t) \tag{4.6}$$

holds, where $f_i = \varphi_i - \psi_i$, $1 \leqslant i \leqslant k$, and $P_2(t)$ is a polynomial of degree 2.

Equation (4.6) can be used to derive results on uniqueness of structural decompositions. For example, suppose that $r = k = 4$, and c.f. of U_i are non-vanishing and have no normal components. Then an application of Lemma 2.3 shows that each f_i is quadratic in t which, in turn, implies that X must have a unique structure of the kind $X = U_1\beta_1 + \ldots + U_4\beta_4$.

The following is perhaps a more interesting result in this direction.

THEOREM 4.1. *Suppose that the random variable X admits the representation*

$$X = U_1\beta_1 + \ldots + U_k\beta_k \tag{4.7}$$

in which the U_i are non-normal variables having no normal components. Assume that the range of $\big((\beta_1, t), 0, \ldots, 0\big)$ is contained in the range of $\big((\beta_1, t), (\beta_2, t), \ldots, (\beta_k, t)\big)$. Then every structure of X admits the representation

$$X = U_1\beta_1 + V_2\beta_2 + \ldots + V_k\beta_k + V_{k+1}\gamma_{k+1} + \ldots + V_l\gamma_l, \tag{4.8}$$

where $V_{k+1}, \ldots, V_l$ are normal variables.

Proof. It follows from Corollary 3 to Lemma 4.1 that every representation of X is of the form

$$X = V_1\beta_1 + \ldots + V_k\beta_k + V_{k+1}\gamma_{k+1} + \ldots + V_l\gamma_l, \tag{4.9}$$

where $V_1, \ldots, V_k$ are non-normal and $V_{k+1}, \ldots, V_l$ are normal variables.

Choose a t such that $(\beta_1, t) = c \neq 0$ and $(\beta_2, t) = \ldots = (\beta_k, t) = 0$. Then from (4.7) and (4.9) we have

$$cU_1 = cV_1 + V_{k+1}(\gamma_{k+1}, t) + \ldots + V_l(\gamma_l, t). \tag{4.10}$$

Since U_1 has no normal component, $(\gamma_{k+1}, t) = \ldots = (\gamma_l, t) = 0$. So $U_1 = V_1$.

COROLLARY 1. *Suppose that, for each i, the range of $\big(0, \ldots, (\beta_i, t), \ldots, 0\big)$ is contained in the range of $\big((\beta_1, t), \ldots, (\beta_i, t), \ldots, (\beta_k, t)\big)$. Then X has a unique structure.*

It is worth-while noting here that the assertion of Corollary 1 also goes through when exactly one of the structural variables in (4.7) has a normal component.

COROLLARY 2. *Suppose that it is known that the U_i in* (4.7) *are non-normal but may have normal components. Then, under the hypotheses of Corollary* 1, X *has a unique structure among all structural decompositions with as many as k variables.*

Theorem 4.1 and its corollaries are extensions of similar results due to Rao (cf. [11], Theorem 10.3.6, p. 312). We refer the reader to Kagan et al. ([11], p. 313-317) for other results of this kind which can be easily extended to Hilbert space-valued random variables.

Motivated by Rao's definition of linear structure, we furnish the following definition of a general linear structure:

Definition 4.2. A Hilbert space-valued random variable X is said to have a *general linear structure* if there exist non-degenerate independent Hilbert space-valued random variables $Y_1, \ldots, Y_k$ and bounded linear operators $A_1, \ldots, A_k$ such that

$$X = \mu + A_1' Y_1 + \ldots + A_k' Y_k, \tag{4.11}$$

where μ is a scalar in H and, for each $i \neq j$, range$(A_i, 0) \subset$ range(A_i, A_j).

Suppose that X has an alternative expression

$$X = \nu + B_1' Z_1 + \ldots + B_l' Z_l. \tag{4.12}$$

Then, choosing a $t \in \bigcap_i \ker(A_i)$ yields $(\mu - \nu, t) = (B_1 t, Z_1) + \ldots + (B_l t, Z_l)$ showing

$$\bigcap_i \ker(A_i) = \bigcap_j \ker(B_j)$$

for any two representations of X.

Further, if ω is an element of the sample space over which the Y and the Z are defined, then

$$\mu + A_1 Y_1(\omega) + \ldots + A_k Y_k(\omega) = \nu + B_1 Z_1(\omega) + \ldots + B_l Z_l(\omega). \tag{4.13}$$

This equation shows that we can (and do) eliminate μ and ν in the representation of X by suitable subtraction.

Based on the above terminology we have the following characterization of normality:

THEOREM 4.2. *Suppose that a random variable X admits the general linear structures*

$$X = A_1' Y_1 + \ldots + A_k' Y_k = B_1' Z_1 + \ldots + B_l' Z_l \tag{4.14}$$

and also suppose that range$(A_1, 0) \subset$ range(A_1, B_j), $1 \leqslant j \leqslant l$. *Then the random variable $A_1' Y_1$ has a normal distribution.*

Proof. Let φ_i and ψ_j denote the l.c.f. of Y_i and of Z_j, respectively, in a neighborhood of the origin. Then, by considering the c.f. of X, we see that φ_i and ψ_j satisfy the following functional equation in a neighborhood of the origin:

$$\varphi_1(A_1 t) + \ldots + \varphi_k(A_k t) - \psi_1(B_1 t) - \ldots - \psi_l(B_l t) = 0.$$

An application of Lemma 2.3 now shows that $\varphi_1(A_1 t)$ is a polynomial of degree less than or equal to $k+l-2$ in a neighborhood of $t = 0$. By a suitable application of Marcinkiewicz's theorem we see that, for each t, $(A_1 t, Y_1) = (t, A_1' Y_1)$ is a normal variable on the line, so $A_1' Y_1$ has a normal distribution.

COROLLARY 1. *Suppose that, for each i,* range$(A_i, 0) \subset$ range(A_i, B_j), $1 \leqslant j \leqslant l$. *Then X has a normal distribution.*

COROLLARY 2. *Let the operators A_i, $1 \leqslant i \leqslant k$, and B_j, $1 \leqslant j \leqslant k$, satisfy the following conditions:* range$(A_i, 0) \subset$ range(A_i, A_j), $i \neq j$, *and* range$(A_i, 0) \subset$ range(A_i, B_j). *Let $Y_1, \ldots, Y_k$ be mutually independent and suppose that $A_1' Y_1 + \ldots + A_k' Y_k$ and $B_1' Y_1 + \ldots + B_k' Y_k$ are identically distributed. Then each $A_i' Y_i$ (so also each $B_j' Y_j$) has a normal distribution.*

Theorem 4.2 and its corollaries can be regarded as suitable generalizations of Lemma 4.1 and its corollaries. Further analogous results concerning general linear structures are under investigation and would be published elsewhere.

5. Characterization of probability laws through distributions of linear functions. In [11] and [23] Rao proved the remarkable result that the probability distribution of $p(p+1)/2$ independent real-valued random variables can be determined, except for a change of location, by the joint probability distribution of p suitable linear functions of the variables. In this section we furnish dimension-free extensions of this result. To illustrate the techniques to be used in this section we consider the following result first:

THEOREM 5.1. *Let X_i, $1 \leqslant i \leqslant 10$, be independent Hilbert space-valued random variables with non-vanishing characteristic functionals. Then the joint probability distribution of the four random variables*

$$\begin{aligned} Y_1 &= X_1 + X_5 + X_6 + X_7, \\ Y_2 &= X_1 + X_5 + X_8 + X_9, \\ Y_3 &= X_1 + X_6 + X_8 + X_{10}, \\ Y_4 &= X_1 + X_7 + X_9 + X_{10} \end{aligned} \tag{5.1}$$

determines the probability distribution of each variable up to a change of location.

Proof. Let φ_i and f_i denote two possible c.f.'s of X_i, $1 \leqslant i \leqslant 10$. Let $\psi_i(t) = \log\varphi_i(t) - \log f_i(A)$. It can then be seen that the ψ_i's satisfy the equation

$$\sum_1^{10} \psi_i(A_i t) = 0, \tag{5.2}$$

where $\boldsymbol{t} = (t_1, t_2, t_3, t_4) \in H^4$ and the A_i are the following row vectors in H^4:

$$A_1 = (1, 0, 0, 0), \quad A_2 = (0, 1, 0, 0), \quad A_3 = (0, 0, 1, 0), \quad A_4 = (0, 0, 0, 1),$$
$$A_5 = (1, 1, 0, 0), \quad A_6 = (1, 0, 1, 0), \quad A_7 = (1, 0, 0, 1), \quad A_8 = (0, 1, 1, 0),$$
$$A_9 = (0, 1, 0, 1), \quad A_{10} = (0, 0, 1, 1).$$

In order to determine ψ_1, say, we note that $A_1A_2' = A_1A_3' = A_1A_4$ $A_1A_8' = A_1A_9' = A_1A_{10}' = 0$ and $\mathrm{range}(A_1, 0, 0, 0) \subset \mathrm{range}(A_1, A_5, A_6, A_7)$. Consequently, by Lemma 2.5, ψ_1 is linear, and so the distribution of X_1 is determined up to a change of location. In a similar manner we show that ψ_2, ψ_3, and ψ_4 are all linear. Therefore, (5.2) determines the probability distribution of X_i, $1 \leqslant i \leqslant 4$, up to a change of location. It thus remains to find the distribution of X_i, $5 \leqslant i \leqslant 10$. To do so, we can (and do) rewrite (5.2) as

$$\sum_{5}^{10} \psi_i(A_i\boldsymbol{t}) = P_1(\boldsymbol{t}), \tag{5.3}$$

where $P_1(\boldsymbol{t})$ is linear in $\boldsymbol{t}$. Now to find the distribution of X_5, for example, note that $A_5A_{10}' = 0$, and $\mathrm{range}(A_5, 0, \ldots, 0) \subset \mathrm{range}(A_5, A_6, A_7, A_8, A_9)$. So, by Lemma 2.5, ψ_5 is linear. In an analogous manner we show that the ψ_i, $5 \leqslant i \leqslant 10$, are all linear. Hence (5.2) determines the distributions of all the X_i up to a change of location.

Theorem 5.1, when extended in an obvious manner, shows that the probability distribution of each of $p(p+1)/2$ random variables can be determined up to a change of location from the joint distribution of p suitable linear functions of the variables. A careful review of numerous characterizations of this kind in the literature shows that Lemma 2.5 plays, as it has in our proof, a central role in establishing such results in the following way: one first reduces the characterization problem to that of solving a functional equation of the kind $\sum\psi_i(A_it) = 0$ in which the ψ_i are polynomials of unknown degree; the best possible degree for ψ_i is then determined by techniques analogous to that of Lemma 2.5.

We present a few additional examples to illustrate these techniques.

THEOREM 5.2. *Let X_1, X_2, X_3 be three independent random variables in R^m with non-vanishing c.f.'s. Let A_1, A_2, A_3 be three given matrices of order $m \times 2m$ each of rank m and let* $\mathrm{rank}(A_1' : A_2') = \mathrm{rank}(A_1' : A_3') = \mathrm{rank}(A_2' : A_3') = 2m$. *Then the probability distribution of $A_1'X_1 + A_2'X_2 + A_3'X_3$ determines the probability distribution of each X_i up to a change of location.*

Proof. Let φ_i and f_i denote two alternative c.f.'s for X_i, $i = 1, 2, 3$. Let $\psi_i = \log\varphi_i - \log f_i$. It is easily seen that the ψ_i satisfy the equation

$$\psi_1(A_1\boldsymbol{t}) + \psi_2(A_2\boldsymbol{t}) + \psi_3(A_3\boldsymbol{t}) = 0, \tag{5.4}$$

where $\boldsymbol{t}$ is a $2m$-vector. By the hypotheses of the theorem, it follows that $\mathrm{range}(A_i, 0) \subset \mathrm{range}(A_i, A_j)$ for each $i \neq j$. So, by Lemma 2.3, each $\psi_i(A_i\cdot)$ is linear. Since the A_i are surjective, each $\psi_i(\cdot)$ is linear. This implies that the c.f. of each X_i is determined up to a change of location.

The following theorem furnishes a similar result for five random variables:

THEOREM 5.3. *Let X_i, $1 \leqslant i \leqslant 5$, be random variables in R^m with non-vanishing c.f.'s. Let A_i, $1 \leqslant i \leqslant 5$, be five $(m \times 3m)$-matrices each of rank m. Suppose that, for each i, there is a permutation (i, i_2, i_3, i_4, i_5) of $(1, 2, 3, 4, 5)$ such that*

$$\mathrm{rank}(A_i' : A_{i_2}' : A_{i_3}') = \mathrm{rank}(A_i' : A_{i_4}' : A_{i_5}') = 3m.$$

Then the probability distribution of $A_1'X_1 + \ldots + A_5'X_5$ determines the probability distribution of each X_i up to a change of location.

Proof. Let $\psi_i = \log\varphi_i - \log f_i$, $1 \leqslant i \leqslant 5$, where φ_i and f_i are two alternative expressions for the c.f. of X_i. It can now be seen that the ψ_i satisfy

$$\sum_1^5 \psi_i(A_i\boldsymbol{t}) = 0, \tag{5.5}$$

where $\boldsymbol{t}$ denotes a $3m$-vector. An application of Lemma 2.4 now shows that each ψ_i is linear which implies the assertion of the theorem.

Theorem 5.1 and Lemma 2.5 motivate the following extension and strengthening of the preceding results:

THEOREM 5.4. *Let X_i, $1 \leqslant i \leqslant n$, be n random variables in a given Hilbert space H with non-vanishing c.f.'s. Let A_i be n bounded surjective linear operators. Suppose that the A_i satisfy the following conditions:*

For each i, $1 \leqslant i \leqslant m$, and $m \leqslant n-3$, there are a $k(i) > i$ and a permutation $(i+1(i), \ldots, n(i))$ of $(i+1, \ldots, n)$ such that $A_iA_{j(i)}' = 0$ for $j = i+1, \ldots, k$ and

$$\mathrm{range}(A_i, 0, \ldots, 0) \subset \mathrm{range}(A_i, A_{k+1(i)}, \ldots, A_{n(i)}).$$

Then the probability distribution of $\sum_1^n A_j'X_j$ determines the distribution of each X_i, $1 \leqslant i \leqslant m$, up to a change of location; and in that case for all

$t \in H$ we have

$$\sum_{m+1}^{n} \psi_j(A_j t) = P_1(t), \tag{5.6}$$

where $P_1(t)$ is linear in t, and $\psi_j(t) = \log\varphi_j(t) - \log f_j(t)$, φ_j and f_j being two alternative c.f.'s of X_j.

Proof. It is easily seen that if the probability distribution of $\sum A_j' X_j$ is given, then the ψ_j must satisfy the equation

$$\sum_{1}^{n} \psi_j(A_j t) = 0. \tag{5.7}$$

Under the hypotheses of the theorem, an application of Lemma 2.5 shows that $\psi_1(t)$ is linear in t. This means that the distribution of $\sum_1^n A_j' X_j$ determines the distribution of X_1 up to a change of location. Also, since $\psi_1(t)$ is linear, we can (and do) rewrite (5.7) as

$$\sum_{2}^{n} \psi_j(A_j t) = P_1(t), \tag{5.8}$$

where $P_1(t)$ is linear in t.

Successive iterations of the above process when applied to (5.8) yield the desired result.

COROLLARY 1. *Let, in Theorem 5.4, $m = n-3$ and* $\text{range}(A_j, 0) \subset \text{range}(A_j, A_k)$ *for $j \neq k = n-2, n-1, n$. Then the distribution of $\sum_1^n A_j' X$ determines the distribution of all the X_i up to a change of location.* (Clear from (5.8) and Lemma 2.3.)

It is worth mentioning at this point that Theorem 5.4 is the outcome of our attempts to strengthen Rao's remarkable result that the joint distribution of as few as p linear functions of $p(p+1)/2$ independent real-valued random variables can determine the distribution of each random variable up to a change of location. Rao's result can be obtained as a special case of ours by letting in Theorem 5.4 $n = p(p+1)/2$ and choosing the A_i to be suitable $(1 \times p)$-matrices as done in Theorem 5.1, for example, in the special case of $n = 10$. It is also worth noting perhaps that many other recent characterization results in vector spaces, analogous to that of Rao (see [6], [12], [14], [15], [19] and [20]), can be easily seen to follow from our theorem. For reasons of brevity, we leave the verification of this last assertion to the reader.

In conclusion we would like to add that results of this paper are by no means complete and do not provide solutions to all the characterization problems that one might encounter. All that we have done here is to furnish general techniques with which these problems can be handled and have barely touched on their applications. We intend to publish a fuller account of our characterization techniques and their applications in a subsequent paper.

REFERENCES

[1] L. V. Ahlfors, *Complex analysis*, New York 1953.

[2] K. L. Chung, *A course in probability theory*, Harcourt 1968.

[3] G. Darmois, *Sur une propriété caractéristique de la lois de Laplace*, Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences 232 (1951), p. 1999-2000.

[4] J. Dieudonné, *Foundations of modern analysis*, New York 1960.

[5] D. Dugue, *Analycité et convexité des fonctions caractéristiques*, Annales de l'Institut Henri Poincaré 12 (1951), p. 45-46.

[6] P. Flusser, *A characterization on the distribution of three independent random variables with values in a locally compact Abelian group*, Sankhyā A 34 (1972), p. 99-110.

[7] S. G. Ghurye and I. Olkin, *A characterization of the multivariate normal distribution*, Annals of Mathematical Statistics 33 (1962), p. 533-541.

[8] V. Goodman and P. K. Pathak, *A dimension free approach to certain characterizations of the normal law*, Sankhyā (1974) (to appear).

[9] — *An extension of a theorem of C. R. Rao*, ibidem (1974) (to appear).

[10] M. H. Ingraham, *Solution of certain functional equations relative to a general linear set*, Transactions of the American Mathematical Society 28 (1926), p. 287-300.

[11] A. M. Kagan, Yu. V. Linnik and C. R. Rao, *Characterization problems in mathematical statistics*, New York 1973.

[12] C. G. Khatri and C. R. Rao, *Functional equations and characterization of probability laws through linear functions of random variables*, Journal of Multivariate Analysis 2 (1972), p. 162-173.

[13] I. I. Kotlarski, *On characterizing the gamma and the normal distribution*, Pacific Journal of Mathematics 20 (1967), p. 69-76.

[14] — *On a characterization of probability distribution by the joint distributions of their linear forms*, Sankhyā A 33 (1971), p. 73-80.

[15] E. Lukacs, *Some recent developments in the theory of multivariate characteristic functions* in *Multivariate analysis II*, Academic Press, 1969, p. 303-319.

[16] Van der Lyn, *La définition fonctionnelle des polynômes dans les groupes abeliens*, Fundamenta Mathematicae 33 (1945), p. 43-50.

[17] S. Mazur und W. Orlicz, *Grundlegende Eigenschaften der polynomischen Operationen*, Studia Mathematica 5 (1934), p. 50-68.

[18] M. A. McKiernan, *On vanishing nth order differences and Hamel bases*, Annales Polonici Mathematici 19 (1967), p. 23-28.

[19] R. P. Pakshirajan and N. R. Mohun, *A characterization of the normal law*, Annals of the Institute of Statistical Mathematics 21 (1969), p. 529-532.

[20] B. L. S. Prakasa Rao, *On a characterization of probability distributions on locally compact Abelian groups*, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 9 (1968), p. 98-100.

[21] B. Ramachandran, *Advanced theory of characteristic functions*, Calcutta 1967.

[22] C. R. Rao, *A decomposition theorem for vector variables with a linear structure*, Annlas of Mathematical Statistics 40 (1969), p. 1845-1849.

[23] — *Characterization of probability laws through linear functions*, Sankhyā A 33 (1971), p. 255-259.

[24] В. П. Скитович, *Линейные формы от независимых случайных величин и нормальный закон распределения*, Известия Академии наук СССР, серия математическая, 18 (1954), p. 185-200.

[25] А. А. Зингер и Ю. В. Линник, *Об одном аналитическом обобщении теоремы Г. Крамера*, Вестник Ленинградского университета 11 (1955), p. 51-56.

Reçu par la Rédaction le 5. 10. 1974