The Annals of Mathematical Statistics
1970, Vol. 41, No. 6, 1889–1895

A NON-LINEAR CHARACTERIZATION OF THE NORMAL DISTRIBUTION

BY ALBERT KINGMAN[1] AND FRANKLIN A. GRAYBILL[2]

Colorado State University

1. Introduction. There are many characterization theorems of the normal distribution that are based on linear combinations of independent, identically distributed (i.i.d.) random variables. For example, some theorems deal with characterizations by one linear function of i.i.d. random variables; some theorems deal with a characterization by the independence of two linear functions of i.i.d. random variables; some theorems deal with characterizations by constant regression. In this paper we shall generalize a characterization first obtained by Skitovich (see [1]) which is as follows.

THEOREM 1.1. *Let $Y_1, Y_2, \cdots, Y_n$ be* i.i.d. *random variables with* cdf *denoted by F and let n be a positive integer. Suppose that L_1 and L_2 are defined by*

$$L_1 = \sum_{i=1}^n a_i Y_i; \qquad L_2 = \sum_{i=1}^n b_i Y_i$$

where the a_i and b_i are constants such that

$$\sum_{i=1}^n a_i b_i = 0 \quad \text{and} \quad \sum_{i=1}^n (a_i b_i)^2 \neq 0.$$

Then F is the cdf *of a normally distributed random variable if and only if L_1 and L_2 are independent.*

We shall state and prove a characterization of the normal distribution that can be considered as an extension of this theorem. We consider the case when the a_i and b_i are *not* constants but are functions of the random variables Y_i.

2. Bivariate case. First we shall consider the case of two random variables Y_1 and Y_2.

THEOREM 2.1. *Assume that Y_1 and Y_2 are* i.i.d. *random variables. Let Y_3 be another random variable which is distributed independently of Y_1 and Y_2 jointly. Let f and g be two measurable functions of the random variable Y_3 such that the functions f and g satisfy*

$$\text{(2.1)} \qquad \text{(i)} \quad f^2(Y_3)+g^2(Y_3) = 1 \quad \text{a.e.}$$

$$\text{(ii)} \quad E[g^{2n-1}(Y_3)] \neq 0 \quad \textit{for} \quad n = 1, 2, 3, \cdots.$$

Define the random variables U and V by

$$U = f(Y_3)Y_1 + g(Y_3)Y_2$$

$$V = g(Y_3)Y_1 - f(Y_3)Y_2.$$

Received October 16, 1969; revised May 15, 1970.

[1] Now at California State College at Hayward.

[2] Research sponsored by NIH Grant T01GM00725 and NSF Grant GP 8835.

Then U and V are independent standard normal random variables if and only if Y_1 and Y_2 are independent standard normal random variables.

PROOF. Suppose that Y_1 and Y_2 are i.i.d. $N(0, 1)$. Let $\phi_{U,V}(t_1, t_2)$ denote the joint characteristic function of the random variables U and V. Then

$$\begin{aligned}
\phi_{U,V}(t_1, t_2) &= E\{\exp(it_1U+it_2V)\} = E\{E[\exp(it_1U+it_2V)\mid Y_3]\} \\
&= E\{E[\exp(it_1f(Y_3)Y_1+it_1g(Y_3)Y_2+it_2g(Y_3)Y_1-it_2f(Y_3)Y_2)\mid Y_3]\} \\
&= E\{E[\exp(i(t_1f(Y_3)+t_2g(Y_3))Y_1+i(t_1g(Y_3)-t_2f(Y_3))Y_2)\mid Y_3]\} \\
&= E\{E[\exp(i(t_1f(Y_3)+t_2g(Y_3))Y_1)\mid Y_3]E[\exp(i(t_1g(Y_3)-t_2f(Y_3))Y_2)\mid Y_3]\} \qquad (2.2)\\
&= E\{\exp[-\tfrac{1}{2}(t_1f(Y_3)+t_2g(Y_3))^2]\cdot\exp[-\tfrac{1}{2}(t_1g(Y_3)-t_2f(Y_3))^2]\} \\
&= E\{\exp[-\tfrac{1}{2}(t_1^2f^2(Y_3)+t_2^2g^2(Y_3)+t_1^2g^2(Y_3)+t_2^2f^2(Y_3))]\} \\
&= \exp(-\tfrac{1}{2}t_1^2-\tfrac{1}{2}t_2^2) \quad \text{for all } t_1 \text{ and } t_2.
\end{aligned}$$

Also, by using a similar argument, we obtain

$$\phi_{U,V}(t_1, 0) = \exp(-\tfrac{1}{2}t_1^2) \tag{2.3}$$

for any t_1, and

$$\phi_{U,V}(0, t_2) = \exp(-\tfrac{1}{2}t_2^2) \tag{2.4}$$

for any t_2.

Hence, for all t_1 and t_2

$$\phi_{U,V}(t_1, t_2) = \phi_{U,V}(t_1, 0)\phi_{U,V}(0, t_2)$$

which is a necessary and sufficient condition for the random variables U and V to be independent. Furthermore, using (2.3) and (2.4), it is easily seen that U and V are each distributed as a standard normal random variable.

Conversely, suppose that U and V are i.i.d. $N(0, 1)$. Then

$$U^2+V^2 = Y_1^2+Y_2^2 \sim \chi^2(2), \tag{2.5}$$

where $\chi^2(k)$ denotes a chi-square distribution with k degrees of freedom, and since Y_1 and Y_2 were assumed i.i.d., it clearly follows that

$$Y_1^2 \sim \chi^2(1). \tag{2.6}$$

That is, the even moments of Y_1 coincide with those of a standard normal random variable. The proof will be complete if it can be shown that

$$E(Y_1^{2n-1}) = 0 \tag{2.7}$$

for all positive integers n, since this sequence of moments uniquely determines the distribution of a standard normal random variable. The proof is by induction on n.

Let $n = 1$ and consider the following expectations.

$$(2.8)\qquad \begin{aligned} E(U) &= E\{f(Y_3)Y_1+g(Y_3)Y_2\} = E\{f(Y_3)\}E\{Y_1\}+E\{g(Y_3)\}E\{Y_2\} \\ &= [E\{f(Y_3)\}+E\{g(Y_3)\}]E\{Y_1\} = 0. \end{aligned}$$

Also

$$(2.9)\qquad \begin{aligned} E(V) &= E\{g(Y_3)Y_1-f(Y_3)Y_2\} = E\{g(Y_3)\}E\{Y_1\}-E\{f(Y_3)\}E\{Y_2\} \\ &= [E\{g(Y_3)\}-E\{f(Y_3)\}]E\{Y_1\} = 0. \end{aligned}$$

It follows that

$$2E\{g(Y_3)\}E\{Y_1\} = 0$$

and hence by condition (ii) of (2.1), that

$$E(Y_1) = 0.$$

Next, let k be some fixed positive integer greater than 1, and assume that for all positive integers $j < k$

$$(2.10)\qquad E\{Y_1^{2j-1}\} = 0.$$

It will then be shown that

$$E\{Y_1^{2k-1})\} = 0.$$

Consider the $(2k-1)$st moment of U.

$$(2.11)\qquad \begin{aligned} E\{U^{2k-1}\} &= E\{\textstyle\sum_{j=0}^{2k-1}\binom{2k-1}{j}[Y_1f(Y_3)]^{2k-j-1}[Y_2g(Y_3)]^j\} \\ &= \textstyle\sum_{j=0}^{2k-1}\binom{2k-1}{j}E\{f^{2k-j-1}(Y_3)g^j(Y_3)\}E\{Y_1^{2k-j-1}\}E\{Y_2^j\} \\ &= E\{f^{2k-1}(Y_3)\}E\{Y_1^{2k-1}\}+E\{g^{2k-1}(Y_3)\}E\{Y_2^{2k-1}\} = 0, \end{aligned}$$

since the odd order moments of U vanish and because of the induction hypothesis (2.10).

Similarly, the $(2k-1)$st moment of V is given by

$$(2.12)\qquad \begin{aligned} E\{V^{2k-1}\} &= E\{\textstyle\sum_{j=0}^{2k-1}\binom{2k-1}{j}[Y_1g(Y_3)]^{2k-j-1}[-Y_2f(Y_3)]^j\} \\ &= \textstyle\sum_{j=0}^{2k-1}(-1)^j\binom{2k-1}{j}E\{g^{2k-j-1}(Y_3)f^j(Y_3)\}E\{Y_1^{2k-j-1}\}E\{Y_2^j\} \\ &= E\{g^{2k-1}(Y_3)\}E\{Y_1^{2k-1}\}-E\{f^{2k-1}(Y_3)\}E\{Y_2^{2k-1}\} = 0. \end{aligned}$$

It follows that

$$(2.13)\qquad 2E\{g^{2k-1}(Y_3)\}E\{Y_1^{2k-1}\} = 0$$

and hence by condition (ii) of the theorem,

$$E\{Y_1^{2k-1}\} = 0$$

and the result in (2.7) is proved and hence the theorem is proved.

EXAMPLE. Suppose that the random variables Y_1, Y_2, Y_3 are i.i.d. such that the random variables

$$U = \frac{Y_1 Y_3}{(Y_3^2 + e^{-2Y_3})^{\frac{1}{2}}} + \frac{Y_2 e^{-Y_3}}{(Y_3^2 + e^{-2Y_3})^{\frac{1}{2}}} \quad \text{and}$$

$$V = \frac{Y_1 e^{-Y_3}}{(Y_3^2 + e^{-2Y_3})^{\frac{1}{2}}} - \frac{Y_2 Y_3}{(Y_3^2 + e^{-2Y_3})^{\frac{1}{2}}}$$

are i.i.d. $N(0, 1)$. Then the random variables Y_1, Y_2, Y_3 are also i.i.d. $N(0, 1)$.

PROOF. The result follows trivially by defining $g(y_3)$ by

$$g(y_3) = \frac{e^{-Y_3}}{(Y_3^2 + e^{-2Y_3})^{\frac{1}{2}}}.$$

3. Multivariate case. The next theorem in one sense generalizes the result obtained in Theorem 2.1, and in another sense is somewhat different from the one obtained there.

Assume that $X_1, X_2, \cdots, X_n, X_{n+1}, X_{n+2}, \cdots, X_{n+m}$ is a set of random variables such that $X_1, X_2, \cdots, X_n$ are i.i.d. Let $\mathbf{y}_1' = (X_1, X_2, \cdots, X_n)$ and let $\mathbf{y}_2' = (X_{n+1}, X_{n+2}, \cdots, X_{n+m})$ denote two vector random variables obtained from the set of random variables $X_1, X_2, \cdots, X_{n+m}$.

THEOREM 3.1. *Suppose that the random variables* $\mathbf{y}_1$ *and* $\mathbf{y}_2$ *are independently distributed and that the component random variables* $X_1, X_2, \cdots, X_n$ *are* i.i.d. *random variables. Let*

$$a_{ij} = f_{ij}(X_{n+1}, X_{n+2}, \cdots, X_{n+m})$$

for $i = 1, 2, \cdots, n$ *and* $j = 1, 2, \cdots, n$ *be Borel measurable functions of the variables* $X_{n+1}, \cdots, X_{n+m}$. *Define the random variables* $U_1, U_2, \cdots, U_n$ *by*

$$\begin{aligned} U_1 &= a_{11}X_1 + a_{12}X_2 + \cdots + a_{1n}X_n \\ U_2 &= a_{21}X_1 + a_{22}X_2 + \cdots + a_{2n}X_n \\ &\vdots \\ U_n &= a_{n1}X_1 + a_{n2}X_2 + \cdots + a_{nn}X_n \end{aligned}$$

which we shall write as

$$\mathbf{U} = \mathbf{A}\mathbf{y}_1$$

where $\mathbf{A} = (a_{ij})$ *is the* $n \times n$ *"random" coefficient matrix. Suppose that the matrix* $\mathbf{A}$ *is orthogonal with probability one and that there exists an index i such that*

$$E[\textstyle\sum_{j=1}^{n} a_{ij}] \neq 0.$$

Then $\mathbf{U} \sim N(\mathbf{0}, \mathbf{I}_n)$ *if and only if* $\mathbf{y}_1 \sim N(\mathbf{0}, \mathbf{I}_n)$.

PROOF. Assume that $\mathbf{y}_1 \sim N(\mathbf{0}, \mathbf{I}_n)$ and let $\phi_U(\mathbf{t})$ denote the joint characteristic function of the random variables $U_1, U_2, \cdots, U_n$. Now

$$\begin{aligned}\phi_U(\mathbf{t}) &= E\{\exp(i\mathbf{t}'\mathbf{U})\} = E\{E[\exp(i\mathbf{t}'\mathbf{U}) \mid \mathbf{y}_2]\} \\ &= E\{E[\exp(i\mathbf{t}'\mathbf{A}\mathbf{y}_1) \mid \mathbf{y}_2]\} = E\{\exp(-\tfrac{1}{2}\mathbf{t}'\mathbf{A}\mathbf{A}'\mathbf{t})\} \\ &= E\{\exp(-\tfrac{1}{2}\mathbf{t}'\mathbf{t})\} = \exp(-\tfrac{1}{2}\mathbf{t}'\mathbf{t}).\end{aligned}$$

That is, $\mathbf{U} \sim N(\mathbf{0}, \mathbf{I}_n)$.

Conversely, assume that $\mathbf{U} \sim N(\mathbf{0}, \mathbf{I}_n)$, and, without loss of generality, assume that

$$E[\textstyle\sum_{j=1}^n a_{1j}] \neq 0. \tag{3.1}$$

Now

$$\mathbf{U}'\mathbf{U} = \mathbf{y}_1'\mathbf{A}'\mathbf{A}\mathbf{y}_1 = \mathbf{y}_1'\mathbf{y}_1 \sim \chi^2(n)$$

by the condition $\mathbf{U} \sim N(\mathbf{0}, \mathbf{I}_n)$. That is,

$$\textstyle\sum_{i=1}^n X_i^2 \sim \chi^2(n). \tag{3.2}$$

Since the X_i's were assumed to be a set of i.i.d. random variables, it follows that

$$X_1^2 \sim \chi^2(1). \tag{3.3}$$

Therefore, the even moments of X_1 coincide with those of a standard normal random variable. The proof will be complete if it can be shown that

$$E(X_1^{2k+1}) = 0 \tag{3.4}$$

for all nonnegative integers k. Mathematical induction will be used to show that (3.4) holds for all nonnegative integers k.

For $k = 0$, one need only compute

$$\begin{aligned}E(U_1) &= E\{\textstyle\sum_{j=1}^n a_{1j}X_j\} = \sum_{j=1}^n E\{a_{1j}X_j\} \\ &= \textstyle\sum_{j=1}^n E\{a_{1j}\}E\{X_j\} = E(X_1)\sum_{j=1}^n E(a_{1j}) = 0\end{aligned}$$

and use (3.1) to obtain that $E(X_1) = 0$. Next, assume that k is a fixed positive integer greater than one, and that for all nonnegative integers $j < k$,

$$E\{X_1^{2j+1}\} = 0. \tag{3.5}$$

Since $\mathbf{U} \sim N(\mathbf{0}, \mathbf{I}_n)$, it follows that, for any set of nonnegative integers $r_1, r_2, \cdots, r_n$,

$$E\{U_1^{r_1}U_2^{r_2} \cdots U_n^{r_n}\} = 0 \tag{3.6}$$

provided at least one r_i is odd. Hence, the expectation

$$E\{U_1(U_1^2 + U_2^2 + \cdots + U_n^2)^k\} = 0.$$

But

$$\begin{aligned}E\{U_1(U_1^2 + \cdots + U_n^2)^k\} &= E\{U_1(X_1^2 + X_2^2 + \cdots + X_n^2)^k\} \\ &= E\{\textstyle\sum_{j=1}^n a_{1j}X_j(X_1^2 + X_2^2 + \cdots + X_n^2)^k\} = 0.\end{aligned} \tag{3.7}$$

Consider the ith term in (3.7).

$$E\{a_{1i}X_i(X_1^2+X_2^2+\cdots+X_n^2)^k\}$$

$$= E\{a_{1i}X_i \textstyle\sum_{r_1+\cdots+r_n=k;r_j\geqq 0} \frac{k!}{r_1!r_2!,\cdots,r_n!}X_1^{2r_1}X_2^{2r_2},\cdots,X_i^{2r_i},\cdots,X_n^{2r_n}\}$$

$$= \textstyle\sum_{r_1+r_2+\cdots+r_n=k;r_j\geqq 0} \frac{k!}{r_1!\cdots r_n!}E\{a_{1i}X_1^{2r_1}\cdots X_i^{2r_i+1}\cdots X_n^{2r_n}\}$$

$$= E\{a_{1i}\} \textstyle\sum_{r_1+\cdots+r_n=k;r_j\geqq 0} \frac{k!}{r_1!\cdots r_n!}E\{X_1^{2r_1}\}\cdots E\{X_i^{2r_i+1}\}\cdots E\{X_n^{2r_n}\}$$

$$= E\{a_{1i}\}E\{X_i^{2k+1}\}, \quad \text{by the induction hypothesis.}$$

Hence, (3.7) becomes

$$\textstyle\sum_{j=1}^{n} E\{a_{1j}\}E\{X_j^{2k+1}\} = E\{X_1^{2k+1}\}E\{\sum_{j=1}^{n} a_{1j}\} = 0$$

and since $E\{\sum_{j=1}^{n} a_{ij}\} \neq 0$, it follows that

$$E\{X_1^{2k+1}\} = 0.$$

Therefore, by mathematical induction,

$$E\{X_1^{2k+1}\} = 0$$

for all nonnegative integers k, and the theorem is proved.

EXAMPLE 3.1. Let $X_1, X_2, X_3, X_4, X_5, X_6$ be i.i.d. random variables. Let

$$\mathbf{a}' = [e^{-X_4}, \ln|X_4X_5|, \sin X_6]$$

and define the matrix $\mathbf{B}$ by

$$\mathbf{B} = \mathbf{aa}'/\mathbf{a}'\mathbf{a}.$$

Hence, the matrix $\mathbf{A}$ defined by

$$\mathbf{A} = \mathbf{I} - 2\mathbf{B}$$

is a symmetric orthogonal matrix with probability one. Now if

$$E\left\{\frac{\sin^2 X_6 + (\ln|X_4X_5|)^2 - e^{-2X_4} - 2e^{-X_4}\ln|X_4X_5| - 2e^{-X_4}\sin X_6}{e^{-2X_4} + (\ln|X_4X_5|)^2 + \sin^2 X_6}\right\} \neq 0$$

then the conditions of the theorem hold for the random variables U_1, U_2, U_3 defined by

$$\mathbf{U} = \mathbf{AX}$$

where $\mathbf{U}' = (U_1, U_2, U_3)$ and $\mathbf{X}' = (X_1, X_2, X_3)$. Thus $\mathbf{U} \sim N(\mathbf{0}, \mathbf{I}_3)$ if and only if $\mathbf{X} \sim N(\mathbf{0}, \mathbf{I}_3)$.

If one compares the assumptions made on the variables $U_1, U_2, \cdots, U_n$ in this example with those given in many of the previously known characterizations, it seems as though the conditions of Theorem 3.1 are much too stringent. This certainly would be the case if the random variables $X_{n+1}, X_{n+2}, \cdots, X_{n+m}$ were always assumed to be dengerate random variables. However, if one allows the a_{ij}'s in the matrix **A** to be non-degenerate random variables (as was done in Theorem 3.1), much stronger assumptions on the distributional properties of the random variables $U_1, U_2, \cdots, U_n$ are needed than those given previously. For example, suppose it had been assumed that the random variables $U_1, \cdots, U_n$ were i.i.d. random variables (not necessarily normally distributed). The following example shows that under these conditions, it is not necessary that X_1 and X_2 should be normally distributed.

EXAMPLE 3.2. Let X_1 and X_2 be i.i.d. random variables with common characteristic function

$$\phi(t) = \cos t$$

for all real t. Suppose that X_3 is a random variable which is independent of the joint distribution of X_1 and X_2 and has the distribution given by

X_3	0	$\frac{1}{2}\pi$
$P(X_3)$	$\frac{1}{2}$	$\frac{1}{2}$

Define U_1 and U_2 by

$$U_1 = X_1 \sin X_3 + X_2 \cos X_3$$

$$U_2 = X_1 \cos X_3 - X_2 \sin X_3.$$

Then the distributions of $\sin X_3$ and $\cos X_3$ are given by

$\sin X_3$	0	1
$P(\sin X_3)$	$\frac{1}{2}$	$\frac{1}{2}$

$\cos X_3$	0	1
$P(\cos X_3)$	$\frac{1}{2}$	$\frac{1}{2}$

Now $E\{\sin X_3 + \cos X_3\} = 1$ and the matrix **A** given by

$$\mathbf{A} = \begin{bmatrix} \sin X_3 & \cos X_3 \\ \cos X_3 & -\sin X_3 \end{bmatrix}$$

is clearly orthogonal with probability one.

It is easily shown that both U_1 and U_2 have the same ch.f. as X_1 and X_2 and that they are independent.

It is also of interest to determine if the condition

$$E\{\textstyle\sum_{j=1}^{n} a_{ij}\} \neq 0 \quad \text{for some} \quad i$$

is required in Theorem 3.1. We have not been able to show this.

REFERENCE

[1] SKITOVICH, V. P. (1962). Linear combinations of independent random variables and the normal distribution law. *Select. Transl. Math. Statist. Prob.* **2** 211–228.