A PROPERTY OF THE NORMAL DISTRIBUTION

By Eugene Lukacs and Edgar P. King[1]

National Bureau of Standards

1. Summary. The following theorem is proved.

Let $X_1, X_2, \cdots, X_n$ be n independently (but not necessarily identically) distributed random variables, and assume that the nth moment of each $X_i (i = 1, 2, \cdots, n)$ exists. The necessary and sufficient conditions for the existence of two statistically independent linear forms $Y_1 = \sum_{s=1}^n a_s X_s$ and $Y_2 = \sum_{s=1}^n b_s X_s$ are:

(A) Each random variable which has a nonzero coefficient in both forms is normally distributed.

(B) $\sum_{s=1}^n a_s b_s \sigma_s^2 = 0$.

Here σ_s^2 denotes the variance of X_s $(s = 1, 2, \cdots, n)$.

For $n = 2$ and $a_1 = b_1 = a_2 = 1$, $b_2 = -1$ this reduces to a theorem of S. Bernstein [1]. Bernstein's paper was not accessible to the authors, whose knowledge of his result was derived from a statement of S. Bernstein's theorem contained in a paper by M. Fréchet [3]. A more general result, not assuming the existence of moments was obtained earlier by M. Kac [4]. A related theorem, assuming equidistribution of the X_i $(i = 1, 2, \cdots n)$ is stated without proof in a recent paper by Yu. V. Linnik [5].

2. Introduction. We consider two linear forms

$$Y_1 = \sum_{s=1}^n a_s X_s, \qquad Y_2 = \sum_{s=1}^n b_s X_s \tag{1}$$

in the n independently distributed random variables $X_1, X_2, \cdots, X_n$. We arrange the variables so that the first p $(X_1, X_2, \cdots, X_p)$ have nonzero coefficients in both forms and the remaining $(n - p)$ have zero coefficients in one form or the other. Clearly $0 \leqq p \leqq n$. When $p = 0$, Y_1 and Y_2 are trivially independent; when $p = 1$, Y_1 and Y_2 cannot be independent. For $p \geqq 2$, it is clear that the statistical independence of the original linear forms (1) is completely equivalent to the independence of the forms $Z_1 = \sum_{s=1}^p a_s X_s$ and $Z_2 = \sum_{s=1}^p b_s X_s$. This means that when $p < n$ the distributions of the random variables $X_{p+1}, \cdots, X_n$ do not affect the independence of Y_1 and Y_2. This is why the theorem contains only a statement about the distributions of those random variables with nonzero coefficients in both forms.

If for some pairs of corresponding coefficients, say the first r $(1 < r < p)$, the relation

$$a_1/b_1 = a_2/b_2 = \cdots = a_r/b_r = C \tag{2}$$

Received 7/6/53.

[1] Now with Eli Lilly and Company

holds, then we can rewrite Z_1 and Z_2 as

$$Z_1 = C(b_1X_1 + \cdots + b_rX_r) + a_{r+1}X_{r+1} + \cdots + a_pX_p ,$$
$$Z_2 = b_1X_1 + \cdots + b_rX_r + b_{r+1}X_{r+1} + \cdots + b_pX_p .$$

Introducing the new variable $\tilde{X}_1 = b_1X_1 + \cdots + b_rX_r$, we see that the independence of Y_1 and Y_2 is equivalent to the independence of the forms $\tilde{Z}_1 = C\tilde{X}_1 + a_{r+1}X_{r+1} + \cdots + a_pX_p$ ant $\tilde{Z}_2 = \tilde{X}_1 + b_{r+1}X_{r+1} + \cdots + b_pX_p$. If the theorem holds for the forms $\tilde{Z}_1$ and $\tilde{Z}_2$, Cramér's theorem [2] shows that the normality of $\tilde{X}_1$ implies the normality of the random variables $X_1, X_2, \cdots, X_r$. We proceed in the same manner if there are several groups of random variables for which a relation of type (2) holds. Hence our problem reduces to the study of the independence of two linear forms whose coefficient matrix contains no vanishing minor of order 2.

Finally it is clear that the independence of Y_1 and Y_2 is equivalent to the independence of the forms $\tilde{Y}_1 = \sum_{s=1}^{n} a_s(X_s - E[X_s])$ and

$$\tilde{Y}_2 = \sum_{s=1}^{n} b_s(X_s - E[X_s]).$$

Therefore we shall assume without loss of generality that the following conditions are satisfied:

$$\text{(i)} \qquad a_sb_s \neq 0, \qquad s = 1, 2, \cdots, n$$

$$\text{(ii)} \qquad a_sb_t - a_tb_s \neq 0 \quad \text{for all } s \neq t;\ s, t = 1, 2, \cdots, n,$$

$$\text{(iii)} \qquad E[X_s] = 0. \qquad s = 1, 2, \cdots, n.$$

3. The functional equation for the characteristic functions. Denote the distribution function of the random variable X_s $(s = 1, \cdots, n)$ by $F_s(x)$ and the corresponding characteristic function by $f_s(t)$. Let $h(u, v)$ be the c.f. of the joint distribution of Y_1 and Y_2 and write $h_1(u) = h(u, 0)$ and $h_2(v) = h(0, v)$. Clearly $h_1(u)$ and $h_2(v)$ are the c.f.'s of the distributions of Y_1 and Y_2, respectively.

We prove first that our conditions are necessary; that is, we assume that Y_1 and Y_2 are statistically independent. In terms of characteristic functions this means

$$h(u, v) = h_1(u)\, h_2(v). \tag{3}$$

Further, because $X_1, \cdots, X_n$ are independent, we have

$$h_1(u) = \prod_{s=1}^{n} f_s(a_s u), \tag{4}$$

$$h_2(v) = \prod_{s=1}^{n} f_s(b_s v), \tag{5}$$

$$h(u, v) = \prod_{s=1}^{n} f_s(a_s u + b_s v). \tag{6}$$

Finally, substituting (4), (5), and (6) in (3) we obtain the following functional equation in the characteristic functions:

$$\prod_{s=1}^{n} f_s(a_s u + b_s v) = \prod_{s=1}^{n} f_s(a_s u) f_s(b_s v). \tag{7}$$

4. The differential equations for the cumulant generating functions. The general procedure for determining the explicit form of the characteristic functions $f_s(t)$ will be to differentiate the logarithm of (7) r times ($r = 1, 2, \cdots, n$) with respect to u, set $u = 0$, and solve the resulting n differential equations for $\ln f_s(t)$ ($s = 1, \cdots, n$).

We first note that $f_s(0) = 1$ ($s = 1, \cdots, n$) and that $f_s(t)$ is a continuous function of t. Therefore there exists a neighborhood of the origin in which all the factors occurring in (7) are different from zero. This neighborhood could of course be the entire plane. In the following derivation we restrict the values of u and v to this neighborhood; then we may take the logarithm of both sides of (7) and obtain

$$\sum_{s=1}^{n} \phi_s(a_s u + b_s v) = \sum_{s=1}^{n} \phi_s(a_s u) + \sum_{s=1}^{n} \phi_s(b_s v), \tag{9}$$

where $\phi_s(x) = \ln f_s(x)$. Differentiating (9) r times with respect to u and setting $u = 0$ yields

$$\sum_{s=1}^{n} \left[\frac{\partial^r}{\partial u^r} \phi_s(a_s u + b_s v)\right]_{u=0} = \sum_{s=1}^{n} \left[\frac{d^r}{du^r} \phi_s(a_s u)\right]_{u=0}. \tag{10}$$

Letting $z_s = a_s u$, we find that the typical term on the left side of (10) becomes

$$\left[\frac{\partial^r}{\partial u^r} \phi_s(a_s u + b_s v)\right]_{u=0} = a_s^r \left[\frac{\partial^r}{\partial z_s^r} \phi_s(z_s + b_s v)\right]_{z_s=0}. \tag{11}$$

With the substitution $\Psi_s(v) = \phi_s(b_s v)$, (11) becomes

$$\left[\frac{\partial^r}{\partial u^r} \phi_s(a_s u + b_s v)\right]_{u=0} = \left(\frac{a_s}{b_s}\right)^r \frac{d^r}{dv^r} \Psi_s(v). \tag{12}$$

Similarly the typical term on the right side of (10) becomes

$$\left[\frac{d^r}{du^r} \phi_s(a_s u)\right]_{u=0} = a_s^r \left[\frac{d^r}{dz^r} \phi_s(z_s)\right]_{z_s=0} = (ia_s)^r \kappa_r^{(s)} \tag{13}$$

where $\kappa_r^{(s)}$ is the rth order cumulant of X_s. Substituting (12) and (13) in (10) we obtain

$$\sum_{s=1}^{n} \xi_s^r \frac{d^r}{dv^r} \Psi_s(v) = \sum_{s=1}^{n} (ia_s)^r \kappa_r^{(s)} \qquad r = 1, 2, \cdots, n \tag{14}$$

where $\xi_s = a_s/b_s$. Differentiating (14) $(n - r)$ times yields the system of differential equations

$$
(15)\qquad \begin{cases} \sum_{s=1}^{n} \xi_s^r \dfrac{d^n}{dv^n} \Psi_s(v) = 0 & r = 1, 2, \cdots, n-1 \\ \sum_{s=1}^{n} \xi_s^n \dfrac{d^n}{dv^n} \Psi_s(v) = \sum_{s=1}^{n} (ia_s)^n \kappa_n^{(s)}. \end{cases}
$$

We have to determine all the distribution functions whose characteristic functions satisfy this system of differential equations and the initial conditions

$$
(15a)\qquad \begin{cases} \left[\dfrac{d^r}{dv^r} \Psi_s(v)\right]_{v=0} = (ib_s)^r \kappa_r^{(s)} & r = 1, 2, \cdots, n-1 \\ \Psi_s(0) = 1. \end{cases}
$$

We now define

$$
D_n = \begin{vmatrix} \xi_1 & \cdots & \cdots & \xi_n \\ \xi_1^2 & \cdots & \cdots & \xi_n^2 \\ \cdots & \cdots & \cdots & \cdots \\ \xi_1^n & \cdots & \cdots & \xi_n^n \end{vmatrix}
$$

and denote by $D_{s,n}$ the cofactor of the element in the sth column and the nth row of D_n. Considering (15) as a system of n linear equations in the quantities $d^n\Psi_s(v)/dv^n$, we obtain the solutions

$$
(16)\qquad \frac{d^n}{dv^n} \Psi_s(v) = \frac{D_{s,n}}{D_n} \sum_{s=1}^{n} (ia_s)^n \kappa_n^{(s)} = i^n C_{s,n}, \quad \text{say.}
$$

Integrating (16) n times and employing the initial conditions (15a) yields

$$
\Psi_s(v) = \sum_{j=1}^{n-1} \frac{(ib_s)^j}{j!} \kappa_j^{(s)} v^j + \frac{C_{s,n}}{n!} (iv)^n.
$$

Since $f_s(b_s v) = \exp[\phi_s(b_s v)] = \exp[\Psi_s(v)]$ we have

$$
(17)\qquad f_s(b_s v) = \exp\left[\sum_{j=1}^{n-1} \frac{\kappa_j^{(s)}}{j!} (ib_s v)^j + \frac{C_{s,n}}{b_s^n n!} (ib_s v)^n\right].
$$

In case any of the functions $f_s(t)$ become zero for some real t, this solution is valid only in a certain neighborhood of the origin. We next show by an indirect proof that none of the functions $f_s(t)$ $(s = 1, \cdots, n)$ has a real zero; from this we can conclude that (17) is valid for all real t.

Let us therefore assume that one or more of the c.f.'s $f_s(t)$ have zeros. In this case at least one of the functions $f_s(b_s v)$ will have a zero. Denote by v_r^0 the zero closest to the origin and by $f_r(t)$ a function for which $f_r(b_r v_r^0) = 0$. For $|v| < |v_r^0|$ we have $f_s(b_s v) \neq 0$ $(s = 1, \cdots, n)$ and formula (17) is valid. Let v be a real number such that $|v| < |v_r^0|$; then we have by (17)

$$
(18)\qquad f_r(b_r v) = \exp\left[\sum_{j=1}^{n-1} \frac{\kappa_j^{(n)}}{j!} (ib_r v)^j + \frac{C_{r,n}}{b_r^n n!} (ib_r v)^n\right].
$$

But $f_r(t)$ is a continuous function. Hence $\lim_{v \to v_r^0} f_r(b_r v) = f_r(b_r v_r^0) = 0$ by assumption. However, from (18) it is clear that

$$\lim_{v \to v_r} f_r(b_r v) = \exp\left[\sum_{j=1}^{n-1} \frac{\kappa_j^{(n)}}{j!} (ib_r v_r^0)^j + \frac{C_{r,n}}{b_r^n n!} (ib_r v_r^0)^n\right]$$

which is always different from zero. This is a contradiction, and hence formula (17) is valid for all values of v. Writing $t = b_s v$ we finally obtain

$$f_s(t) = \exp\left[\sum_{j=1}^{n-1} \frac{\kappa_j^{(s)}}{j!} (it)^j + \frac{C_{s,n}}{b_s^n n!} (it)^n\right]. \tag{19}$$

5. Proof of the theorem. We have determined all the solutions of the system (15) satisfying the initial conditions (15a). In order to find the distribution functions whose characteristic functions satisfy this system we must select those functions (19) which are characteristic functions. This is easily done by means of the following result [6].

Theorem of Marcinkiewicz. No function of the form $e^{a_0 + a_1 z + \cdots + a_r z^r}$ $(r > 2)$ *can be a characteristic function.*

Hence the degree of the polynomial in (19) cannot exceed 2. In case $n > 2$ we must have

$$\kappa_j^{(s)} = 0 \qquad j = 3, 4, \cdots, n-1; s = 1, 2, \cdots, n; n > 3$$

$$C_{s,n} = 0 \qquad s = 1, 2, \cdots, n; n > 2.$$

Because the factor $D_{s,n}/D_n$ in $C_{s,n}$ cannot vanish, these relations reduce to

$$\begin{cases} \kappa_j^{(s)} = 0 & j = 3, \cdots, n-1; \quad n > 3 \\ \sum_{s=1}^{n} a_s^n \kappa_n^{(s)} = 0 & n > 2. \end{cases} \tag{20}$$

There is no restriction if $n = 2$. In view of (iii), $\kappa_1^{(s)} = 0$ also, and (19) becomes

$$f_s(t) = \exp[-\tfrac{1}{2}\sigma_s^2 t^2] \qquad n > 2. \tag{21}$$

This shows that each X_s $(s = 1, \cdots, n)$ must be normally distributed, which is condition (A) of the theorem. All cumulants of order $r > 2$ vanish for a normal distribution, hence equations (20) impose no additional restrictions. In case $n = 2$ we have

$$f_s(t) = \exp[-\tfrac{1}{2}kt^2] \qquad n = 2, \tag{22}$$

where k is determined from (16) and (19). The independence of Y_1 and Y_2 implies that they are uncorrelated which yields condition (B) and completes the first part of the proof.

It is easy to establish that conditions (A) and (B) are also sufficient. Assuming that (A) and (B) hold, it follows that Y_1 and Y_2 are uncorrelated and normally distributed. Hence Y_1 and Y_2 must be independent.

For $n = 2$ and $a_1 = a_2 = b_1 = 1$, $b_2 = -1$ we obtain from (22)

$$f_s(t) = \exp\left[-(\sigma_1^2 + \sigma_2^2)t^2/2\right] \qquad s = 1, 2.$$

This shows that $\sigma_1^2 = \sigma_2^2$ and establishes Bernstein's theorem.

REFERENCES

[1] S. BERNSTEIN, "Sur une proprieté caractéristique de la loi de Gauss, *Trans. Leningrad Polytechnic Institute* (1941).

[2] H. CRAMÉR, "Über Eine Eigenschaft Der Normalen Verteilungsfunktion," *Math. Z.* Vol. 41 (1936), pp. 405–414.

[3] M. FRÉCHET, "Généralisation de la loi de probabilité de Laplace, *Ann. Inst. H. Poincaré*, Vol. 13 (1951), pp. 1–29.

[4] M. KAC, "A characterization of the normal distribution," *Amer. J. Math.*, Vol. 61 (1939), pp. 726–728.

[5] YU. V. LINNIK, "Remarks concerning the classical derivation of Maxwell's law," *Doklady Akad. Nauk SSSR*, Vol. 85 (1952), pp. 1251–1254.

[6] J. MARCINKIEWICZ, "Sur une proprieté de la loi de Gauss," *Math. Z.* Vol. 44 (1938), pp. 612–618.

ADDENDUM

The authors are indebted to Professor G. Darmois for calling their attention to his note in the *C. R. Acad. Sci. Paris*, Vol. 232 (1951), pp. 1999–2000 in which he proved the theorem for $n = 2$ without assuming the existence of moments. He later extended this to the case of arbitrary n. His paper will be published in the *Bulletin of the International Statistical Institute*. The method of proof used by Professor Darmois is different from the one presented in this paper. The authors learned that these results were also obtained by methods similar to Darmois' by B. V. Gnedenko (*Izvestiya Akad. Nauk. SSSR, Ser. Mat.*, Vol. 12 (1948), pp. 97–100) for the case $n = 2$ and by V. P. Skitovich (*Doklady Akad. Nauk. SSSR* (N.S.) Vol. 89 (1953), pp. 217–219) for any n.

While reading the proofs of this paper the authors learned that the theorem was also discussed by M. Loève in the appendix to P. Lévy's "Processus stochastiques", Gauthier-Villars, Paris, 1948, pp. 337–338.

ON OPTIMAL SYSTEMS[1]

BY DAVID BLACKWELL

Howard University

1. Summary. For any sequence $x_1, x_2, \cdots$ of chance variables satisfying $|x_n| \leqq 1$ and $E(x_n \mid x_1, \cdots, x_{n-1}) \leqq -u(\max |x_n| \mid x_1, \cdots, x_{n-1})$, where u is a fixed constant, $0 < u < 1$, and for any positive number t,

$$\Pr\{\sup_n (x_1 + \cdots + x_n) \geqq t\} \leqq \left(\frac{1-u}{1+u}\right)^t.$$

Received 7/14/53.

[1] This paper was written under contract Nonr-1197(00) with the Office of Naval Research.