CHARACTERIZATION OF THE NORMAL DISTRIBUTION II

BY RYOICHI SHIMIZU

(Received October 3, 1962)

1. Introduction

Let $x_1, \cdots, x_n$ be a random sample from a certain population with distribution function $F(x)$, and let $C=(c_{ij})$ be an $n\times n$ orthogonal matrix. Define $y_1, \cdots, y_n$ by

$$y_i=\sum_{j=1}^{n} c_{ij}x_j \qquad i=1, 2, \cdots, n.$$

It is well known that $F(x)$ is a normal distribution with zero mean if and only if, for every orthogonal matrix C,

(i) $y_1, y_2, \cdots, y_n$ are independent, and

(ii) every y_i is distributed according to $F(x)$.

But the normality of $F(x)$ is a consequence of a much weaker condition than (i) as well as than (ii). G. Darmois and Skitovich* independently proved the following:

THEOREM. *Let $x_1, \cdots, x_n$ be independently (but not necessarily identically) distributed random variables. If there exist non-zero constants $a_1, a_2, \cdots, a_n, b_1, \cdots, b_n$ such that, $a_1x_1+\cdots+a_nx_n$ and $b_1x_1+\cdots+b_nx_n$ are independent, then each x_i is normally distributed.*

It is clear that the assumption of this theorem is weaker than (i) and that it is independent of (ii).

The purpose of this paper is to show that the normality also follows from an assumption which is weaker than (ii), if we can assume the existence of finite variance. The extension to the multivariate case and to the symmetric stable distribution is also stated.

In what follows, we always assume that $n\geqq 2$ and $x_1, \cdots, x_n$ (or $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$) are independent and identically distributed non-degenerate uni-variate (or p-variate) random variables (random colume vectors) with the distribution function $F(x)$ (or $F(\boldsymbol{x})$). $\varphi(t)$ (or $\varphi(\boldsymbol{t})$, $\boldsymbol{t}=(t_1, \cdots, t_p)$) denotes its characteristic function and $-f(t)=\log\varphi(t)$ (or $-f(\boldsymbol{t})=\log\varphi(\boldsymbol{t})$) the cumulant generating function.

If, for univariate case, $F(x)$ has finite variance σ^2, then $\varphi(t)$ and $f(t)$ have continuous second derivatives and,

$$f''(0)=-\varphi''(0)=\sigma^2.$$

* See, e.g., [2].

Similary, for p-variate case, if $F(\boldsymbol{x})$ has finite second moments, and if Σ is its variance-covariance matrix, then $\varphi(\boldsymbol{t})$ and $f(\boldsymbol{t})$ have continuous partial second derivatives, and

$$f(0)=\left(\frac{\partial^2 f(t_1,\cdots,t_p)}{\partial t_i \partial t_j}\right]_{t=0}\Bigg)_{i,j=1,\cdots,p}=\Sigma.$$

2. Univariate normal distribution $N(\mu, \sigma^2)$

We first prove the following:

THEOREM 1. *Let $x_1, x_2, \cdots, x_n$ be a random sample from $F(x)$ with mean μ and finite variance σ^2. If there exist non-zero constants $a_1, \cdots, a_n$ such that $a_1x_1+\cdots+a_nx_n$ is distributed according to $F(x)$, then the distribution is normal $N(\mu, \sigma^2)$.*

PROOF. In order to avoid unnecessary complexity, we prove the theorem only for $n=2$. The general case will be proved by the slight notational modifications. We write a and b instead of a_1 and a_2. By the assumption we have,

$$f(t)=f(at)+f(bt), \tag{2.1}$$

and by differentiating,

$$f'(t)=af'(at)+bf'(bt) \tag{2.2}$$

$$f''(t)=a^2f''(at)+b^2f''(bt). \tag{2.3}$$

From (2.3), we have, for every positive integer m,

$$f''(t)=\sum_{k=0}^{m}\binom{m}{k}(a^2)^k(b^2)^{m-k}f''(a^kb^{m-k}t). \tag{2.4}$$

Putting $t=0$, we have

$$\sigma^2=\sum_{k=0}^{m}\binom{m}{k}(a^2)^k(b^2)^{m-k}\sigma^2. \tag{2.5}$$

From (2.4) and (2.5),

$$\begin{aligned}|f''(t)-\sigma^2|&=\left|\sum_{k=0}^{m}\binom{m}{k}(a^2)^k(b^2)^{m-k}(f''(a^kb^{m-k}t)-\sigma^2)\right|\\&\leqq\sum_{k=0}^{m}\binom{m}{k}(a^2)^k(b^2)^{m-k}|f''(a^kb^{m-k}t)-\sigma^2|.\end{aligned}$$

Since, from (2.5), $0<|a|, |b|<1$, and since $f''(t)$ is continuous at $t=0$, we can make

$$|f''(a^kb^{m-k}t)-\sigma^2|$$

smaller than any given $\delta>0$, by choosing, m sufficiently large (depending on t). Hence we have

$$|f''(t)-\sigma^2|\leqq\delta\sum_{k=0}^{m}\binom{m}{k}(a^2)^k(b^2)^{m-k}=\delta.$$

This implies $f''(t)=\sigma^2$, from which we get

$$\varphi(t)=\exp\left\{i\mu t-\frac{1}{2}\sigma^2t^2\right\} \tag{2.6}$$

This is the characteristic function of the normal distribution $N(\mu, \sigma^2)$. q.e.d.

Remark. Cumulant generating function $f(t)$ is, in general, defined only in the suitable neighbourhood N of the origin. But, in view of the fact that $0<|a|, |b|<1$, (2.4), (2.5) and hence (2.6) hold in N.

3. Symmetric stable distributions

Let x_1, and x_2 be a random sample from $F(x)$. If, to every pair of positive constants a and b, there corresponds a positive constant c, such that ax_1+bx_2 is distributed as $F(cx)$, then the distribution is called a stable distribution. The characteristic function of the stable distribution can be expressed as

$$\varphi(t)=\exp\left(-c_0+ic_1\frac{t}{|t|}\right)|t|^\alpha$$

where

$$0<\alpha\leqq 2 \qquad 0<c_0$$

$$\left|c_1\cos\frac{\pi\alpha}{2}\right|\leqq c_0\sin\frac{\pi\alpha}{2}.$$

If $c_1=0$, the distribution is symmetric about 0. α is referred to as "the characteristic exponent" of the distribution.

Normal distribution $N(0, \sigma^2)$ is the stable distribution with characteristic exponent $\alpha=2$. It is also characterized as the stable distribution with finite variance.

The following theorem states that "If to every pair of positive constants a and b, ⋯" in the above definition of the stable distribution can be replaced by "If to some a and b ⋯ ."

THEOREM 2. *Let the cumulant generating function be such that, for some constant* α $(2\geqq\alpha>0)$,

$$\frac{f(t)}{|t|^\alpha}\to c_0(\neq 0) \qquad \text{as } t\to 0.$$

If there exist non-zero constants a and b for which ax_1+bx_2 *is distributed as* $F(x)$, *then the distribution is the symmetric stable distribution with characteristic exponent* α.

PROOF. From the assumption, we have

$$f(t)=f(at)+f(bt)\ . \tag{3.1}$$

Hence, for any non-zero t, and positive integer m,

$$\frac{f(t)}{|t|^{\alpha}}=|a|^{\alpha}\frac{f(at)}{|at|^{\alpha}}+|b|^{\alpha}\frac{f(bt)}{|bt|^{\alpha}} \tag{3.2}$$

$$=\sum_{k=0}^{m}\binom{m}{k}(|a|^{\alpha})^{k}(|b|^{\alpha})^{m-k}\frac{f(a^{k}b^{m-k}t)}{|a^{k}b^{m-k}t|^{\alpha}}\ . \tag{3.3}$$

Letting $t\to 0$ in (3.3), we have

$$c_0=\sum_{k=0}^{m}\binom{m}{k}(|a|^{\alpha})^{k}(|b|^{\alpha})^{m-k}c_0\ . \tag{3.4}$$

Using the analogous discussion as in the proof of theorem 1, we get

$$\frac{f(t)}{|t|^{\alpha}}=c_0 \qquad (>0)$$

and

$$\varphi(t)=\exp(-c_0|t|^{\alpha}). \qquad \text{q.e.d.}$$

Remark. The assumption on $f(t)$ may be replaced by the moment condition, e.g.,

$$E(|x|^{\delta})<\infty \qquad 0\leqq\delta<\alpha.$$

But the author has not succeeded in proving this.

4. Multivariate normal distribution $N(\boldsymbol{\mu}, \Sigma)$

The result of theorem 1 can be extended in natural way to the multivariate case.

When the cumulant generating function $f(\boldsymbol{t})$ has continuous second partial derivatives (or equivalently, when $F(\boldsymbol{x})$ has the second moments), we define

$$\boldsymbol{f}(\boldsymbol{t})=\left(\frac{\partial^2 f(t_1,\cdots,t_p)}{\partial t_i\partial t_j}\right)_{i,j=1,\cdots,p}\ .$$

THEOREM. 3. *Let* $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n$ *be a random sample from* $F(\boldsymbol{x})$ *with finite second moments, and let* $\boldsymbol{\mu}$ *and* Σ *be its mean vector and variance-covariance matrix respectively. If there exist non-singular* $p\times p$ *matrices*

$A_1, \cdots, A_n$ for which $A_1x_1 + \cdots + A_nx_n$ is distributed according to $F(x)$, then the distribution $F(x)$ is p-variate normal $N(\mu, \Sigma)$.

PROOF. For simplicity, we prove the theorem only for $n=2$, and write A and B instead of A_1 and A_2. Since Σ is positive definite, there exists a non-singular matrix C such that, $C\Sigma C' = I$. Clearly, the problem is unchanged, if we replace x_i by Cx_i, of which the moment matrix is I. Hence, we assume without any loss of generality that $\Sigma = I$. Now, by the assumption of the theorem, we have

$$f(t) = f(tA) + f(tB) \tag{4.1}$$

and, for every positive integer m,

$$f(t) = Af(tA)A' + Bf(tB)B' \tag{4.2}$$

$$= \sum_k C_{m,k} f(tC_{m,k}) C'_{m,k} \tag{4.3}$$

where
$$C_{m,k} = C_1^{(k)} \cdots C_m^{(k)},$$
$$C_i^{(k)} = A \text{ or } B,$$

and the summation extends over all 2^m terms of the form

$$C_{m,k} = A^{\lambda_1} B^{\mu_1} A^{\lambda_2} B^{\mu_2} \cdots A^{\lambda_m} B^{\mu_m},$$

where
$$\Sigma\lambda_i + \Sigma\mu_i = m, \qquad \lambda_i \geqq 0, \ \mu_i \geqq 0.$$

Putting $t=0$ in (4.2) and (4.3), we have

$$I = AA' + BB' \tag{4.4}$$

$$= \sum_k C_{m,k} C'_{m,k}\,. \tag{4.5}$$

From (4.4) and non-singularity of A and B, we see that

$$\rho^2 \equiv \text{maximum eigenvalue of } AA' \text{ and } BB' < 1.$$

Hence,

$$\|tC_{m,k}\|^2 = \|tC_1^{(k)} \cdots C_m^{(k)}\|^2 \leqq \rho^{2m} \|t\|^2 \to 0 \text{ as } m \to \infty.$$

Continuity of $f(t)$ together with (4.3) and (4.5) implies $f(t) = I$, from which, $\varphi(t) = \exp(it\mu' - tt'/2)$. q.e.d.

Acknowledgement

Thanks are due to Mr. M. Sibuya and Mr. K. Isii for their useful advices.

THE INSTITUTE OF STATISTICAL MATHEMATICS

REFERENCES

[1] H. Cramér, *Mathematical Method of Statistics*, Princeton, 1946.
[2] R. G. Laha, "On a characterization of the normal distribution from properties of suitable linear statistics," *Ann. Math. Stat.* Vol. 28 (1957), pp. 126–139.
[3] P. Lévy, *Théorie de l'addition des variables aléatoires*, Gauthier-Villars, Paris, 1937.

Added in proof

Yu, V. Linnik, "Linear statistics and the normal law, Dokl. Akad. Nauk SSSR. 83 (1952), 353–355", obtained necessary and sufficient condition for the equivalence of the two assertions: (A) the population is normal and (B) two linear statistics are identically distributed. The author read his paper in English translation in "Selected Translations in Math. Stat. and Prob., Inst. of Math. Stat. and Amer. Math, Soc. Vol. (1961)", in which his result is stated without proof. Though his problem is somewhat different from that of mine, the result of theorem 1 seems to be derived from Linnik's result.