ON A CHARACTERIZATION OF THE STABLE LAW WITH FINITE EXPECTATION

By R. G. Laha

Indian Statistical Institute

1. Introduction and summary. A remarkable characterization of the normal law is that if x and y are two independent chance variables such that two linear functions, $ax + by$ $(ab \neq 0)$ and $cx + dy$ $(cd \neq 0)$, are distributed independently of each other, then both x and y are normally distributed. This theorem has been proved without any assumption about the existence of moments by Darmois [2], extending earlier results of Gnedenko [4] and Kac [5]. The question that naturally arises is how far the condition of stochastic independence is necessary, or, in other words, whether the above theorem can be generalised after relaxing the condition of stochastic independence of the linear functions of two independent chance variables. But it is evident that we can always construct two linear functions of non-normal mutually independent chance variables such that they are not independent in the probability sense. In the present paper we shall investigate the nature of the distribution law that may be obtained by imposing the mild restriction of the linearity of regression of one linear function on the other, which is, of course, weaker than the assumption of stochastic independence. We shall prove a general theorem from which a number of results will follow as special cases. But it should be noted that the statements regarding regression or conditional expectation require the assumption that the conditional distribution function exist, and in the following, this assumption will be tacitly made wherever needed.

2. Results. First of all, we shall give a short proof of the following lemma of Rao [8], Rothschild and Mourier [10].

Lemma. *Let x and y be two proper random variables each having a finite expectation (which we may assume to be zero without any loss of generality) such that the regression of y on x exists. Then the necessary and sufficient condition for the regression of y on x to be linear is that*

$$\left.\frac{\partial\varphi(u, v)}{\partial v}\right]_{v=0} = \beta \frac{d\varphi(u, 0)}{du},$$

where $\varphi(u, v)$ stands for the characteristic function of the joint cumulative distribution of x and y, and β is a constant.

Proof of Necessity. Since $\varphi(u, v)$ represents the characteristic function of

Received October 11, 1954; revised March 17, 1955.

the joint cumulative distribution of x and y, we have

$$\begin{aligned}\varphi(u, v) &= E(e^{iux+ivy}) \\ &= \int e^{iux+ivy}\, dF(x, y) \\ &= \int e^{iux}\left[\int e^{ivy}\, dF_x(y)\right] dF(x),\end{aligned}$$

where $F_x(y)$ represents the conditional distribution function of y for fixed x.

But since the expectations of both x and y are assumed to exist and to be equal to zero and, further, the regression of y on x is linear, we must have

$$\begin{aligned}\left.\frac{\partial\varphi(u, v)}{\partial v}\right]_{v=0} &= i\int e^{iux}\left[\int y\, dF_x(y)\right] dF(x) \\ &= i\beta\int e^{iux}x\, dF(x) \\ &= \beta\frac{d\varphi(u, 0)}{du}.\end{aligned}$$

PROOF OF SUFFICIENCY. Since the regression of y on x is assumed to exist, let us denote it by $E_x(y)$, so that $E_x(y) = \int y\, dF_x(y)$.

Then proceeding as above, it can be very easily shown that

$$\left.\frac{\partial\varphi(u, v)}{\partial v}\right]_{v=0} = i\int e^{iux}[E_x(y)]\, dF(x).$$

Hence, the condition

$$\left.\frac{\partial\varphi(u, v)}{\partial v}\right]_{v=0} = \beta\frac{d\varphi(u, 0)}{du}$$

gives

$$\int e^{iux}\,[E_x(y) - \beta x]\, dF(x) = 0.$$

Then, from the uniqueness theorem of Fourier transforms of functions of bounded variation, it follows that $E_x(y) \equiv \beta x$, for all x, except for a set of probability measure zero.

THEOREM 1. *Let x, ξ, and η be three proper random variables, each having a finite expectation (which may be assumed to be zero without any loss of generality) such that x is distributed independently of the joint distribution of ξ and η, but ξ and η have a joint distribution where the regression of η on ξ exists and is linear and given by $E_\xi(\eta) = \beta_0\xi$. Then the regression of $Y = cx + \eta$ on $X = ax + \xi$, $(a \neq 0)$, is always linear irrespective of the distribution functions of x, ξ, and η, whenever the relationship $c = a\beta_0$ is satisfied.*

PROOF. Let $\Phi(u, v)$, $\varphi(u, v)$, and $\varphi_1(u)$ represent the characteristic functions of the joint cumulative distribution of (X, Y), (ξ, η), and the cumulative distribution of x, respectively. Then,

$$\begin{aligned} \Phi(u, v) &= E\{e^{iuX+ivY}\} \\ &= E\{e^{iu(ax+\xi)+iv(cx+\eta)}\} \qquad (1) \\ &= \varphi_1(au + cv)\varphi(u, v). \end{aligned}$$

Now, differentiating both sides of (1) with respect to v and then putting $v = 0$, we get

$$\left.\frac{\partial\Phi(u, v)}{\partial v}\right]_{v=0} = c\varphi_1'(au)\varphi(u, 0) + \left.\frac{\partial\varphi(u, v)}{\partial v}\right]_{v=0} \varphi_1(au). \qquad (2)$$

But, using the lemma above, since $E_\xi(\eta) = \beta_0\xi$,

$$\left.\frac{\partial\varphi(u, v)}{\partial v}\right]_{v=0} = \beta_0 \frac{d\varphi(u, 0)}{du}. \qquad (3)$$

Next, substituting (3) in (2), we get

$$\left.\frac{\partial\Phi(u, v)}{\partial v}\right]_{v=0} = c\varphi_1'(au)\varphi(u, 0) + \beta_0\varphi'(u, 0)\varphi_1(au). \qquad (4)$$

Again, putting $v = 0$ in (1) and then differentiating both sides with respect to u, we get

$$\frac{d\Phi(u, 0)}{du} = a\varphi_1'(au)\varphi(u, 0) + \varphi'(u, 0)\varphi_1(au). \qquad (5)$$

Now, if $c = a\beta_0$, substituting this value of c in (4) and then comparing with (5), we get easily

$$\left.\frac{\partial\Phi(u, v)}{\partial v}\right]_{v=0} = \beta_0 \frac{d\Phi(u, 0)}{du}. \qquad (6)$$

Then from the lemma above, it follows that the regression of Y on X is always linear, whatever may be the distribution function of x, ξ, and η. From Theorem 1, it follows that if η and ξ are stochastically independent, and further if $c = 0$, then the regression of Y on X is always linear, since in this case $\beta_0 = 0$ and the relationship $c = a\beta_0$ is satisfied.

Similarly, if $\xi = by$ $(b \neq 0)$ and $\eta = dy$ $(d \neq 0)$, and further if $bc = ad$, the regression of Y on X is always linear.

3. Further results.

THEOREM 2. *With the same notations and assumptions as used in Theorem* 1, *the necessary and sufficient condition for the regression of Y on X to be linear for all a contained in a closed interval (a_1, a_2), where either $a_1 < a_2 < 0$ or $0 < a_1 < a_2$, and for some c for which the relationship $c \neq a\beta_0$ is satisfied for all a in the interval, is that both x and ξ should belong to a class of stable law with finite expectation.*

PROOF OF NECESSITY. Using the above lemma, the condition of the linearity of regression Y on X gives the relation,

$$\left.\frac{\partial\Phi(u, v)}{\partial v}\right]_{v=0} = \beta\frac{d\Phi(u, 0)}{du}. \tag{7}$$

Next, using (4), (5), and (7) together, we get, after a little rearrangement of terms,

$$(c - a\beta)\varphi_1'(au)\varphi(u, 0) = (\beta - \beta_0)\varphi_1(au)\varphi'(u, 0). \tag{8}$$

We shall first show that in (8) neither $c - a\beta$ nor $\beta - \beta_0$ can be equal to zero under the conditions of the theorem.

Let us suppose that $\beta - \beta_0 = 0$ when $c - a\beta \neq 0$. In this case, (8) reduces to

$$\varphi_1'(au)\varphi(u, 0) = 0. \tag{9}$$

Since $\varphi(u, 0)$ is continuous and equal to unity at the origin, $u = 0$, there always exists a neighbourhood, say $u_0 > 0$, such that for all $|u| < u_0$, we have $\varphi(u, 0) \neq 0$. Then it follows from (9) that for $|u| < u_0/a_0$, we have

$$\varphi_1'(u) = 0, \tag{10}$$

where a_0 is the larger of $|a_1|$ and $|a_2|$. This implies that the distribution of x itself is improper, the whole mass being concentrated at the origin $x = 0$.

Similarly, if $c - a\beta = 0$ when $\beta - \beta_0 \neq 0$, (8) reduces to

$$\varphi_1(au)\varphi'(u, 0) = 0. \tag{11}$$

From (11), proceeding exactly as above, it can be shown that the distribution of ξ itself is improper, the whole mass being concentrated at the origin, $\xi = 0$. But both these cases contradict the conditions of the theorem. Now the only alternative left is when both $c - a\beta$ and $\beta - \beta_0$ vanish simultaneously. But in this case we have $c = a\beta_0$, which is again contrary to the conditions of the theorem.

Now it may be noted that both $\varphi_1(u)$ and $\varphi(u, 0)$ may have real roots. Let ϵ and δ denote the smallest of the absolute values of the real roots of $\varphi_1(u)$ and $\varphi(u, 0)$, respectively. Since both $\varphi_1(u)$ and $\varphi(u, 0)$ are continuous functions of u and since $\varphi_1(0) = \varphi(0, 0) = 1$, it follows that $\epsilon > 0$ and $\delta > 0$. Then, restricting the values of a to an interval I, (a_1, a_2), for which $|a| < \epsilon/\delta$, we can always take the neighbourhood of the origin to be defined by $|u| < \delta$. Thus we have proved the existence of a neighbourhood $|u| < \delta$ of the origin and of an interval I, (a_1, a_2), such that both $\varphi_1(u)$ and $\varphi(u, 0)$ do not vanish if $a \varepsilon I$ and $|u| < \delta$.

Then, confining the values of u and a in these intervals, since the product $\varphi_1(au)\varphi(u, 0) \neq 0$, we may divide both sides of (8) by $\varphi_1(au)\varphi(u, 0)$ and thereby obtain

$$(c - a\beta)\frac{\varphi_1'(au)}{\varphi_1(au)} = (\beta - \beta_0)\frac{\varphi'(u, 0)}{\varphi(u, 0)}. \tag{12}$$

Next, integrating (12) with respect to u, we get

$$\ln \varphi_1(au) = \frac{a(\beta - \beta_0)}{c - a\beta} \ln \varphi(u, 0), \tag{13}$$

where the constant of integration vanishes by virtue of the fact that $\ln \varphi_1(0) = \ln \varphi(0, 0) = 0$.

Since the first moment of x exists, $\ln \varphi_1(au)$ is differentiable with respect to a in the interval (a_1, a_2). Thus it follows that $\theta(a) = a(\beta - \beta_0)/(c - a\beta)$ must also be differentiable with respect to a in the same interval; denoting this derivative by $\theta'(a)$, we may write

$$u \frac{\varphi_1'(au)}{\varphi_1(au)} = \theta'(a) \ln \varphi(u, 0). \tag{14}$$

Again, from the conditions of the theorem, $\theta(a) \neq 0$ for all a in the interval (a_1, a_2). Hence, using (12) and (14) together, we get

$$\begin{aligned} u \frac{\varphi'(u, 0)}{\varphi(u, 0)} &= a \frac{\theta'(a)}{\theta(a)} \ln \varphi(u, 0) \\ &= \lambda \ln \varphi(u, 0), \end{aligned} \tag{15}$$

where $\lambda = a\theta'(a)/\theta(a)$ for all a contained in the interval (a_1, a_2), and thus it follows evidently from (15) that λ is independent of a.

Then, excluding the origin from the interval $|u| < \delta$, that is, in the intervals $(0 < u < +\delta)$ and $(-\delta < u < 0)$, we may divide both sides of (15) by

$$u \ln \varphi(u, 0),$$

and obtain

$$\frac{1}{\ln \varphi(u, 0)} \cdot \frac{\varphi'(u, 0)}{\varphi(u, 0)} = \lambda \frac{1}{u}. \tag{16}$$

Hence, integrating (16) with respect to u, we get

$$\ln \ln \varphi(u, 0) = \begin{cases} \lambda \log |u| + \log c_1, & \text{for } 0 < u < +\delta \\ \lambda \log |u| + \log c_2, & \text{for } -\delta < u < 0. \end{cases} \tag{17}$$

Now, (17) evidently leads to the relation

$$\varphi(u, 0) = \begin{cases} e^{c_1 |u|^\lambda}, & \text{for } 0 < u < +\delta \\ e^{c_2 |u|^\lambda}, & \text{for } -\delta < u < 0, \end{cases} \tag{18}$$

where c_1 and c_2 are the constants of integration. But it is well known that necessary conditions for a function $\varphi(t)$ to be a characteristic function are:

(i) $\varphi(0) = 1$, (ii) $|\varphi(t)| \leqq 1$, and (iii) $\varphi(-t) = \overline{\varphi(t)}$.

Hence, it evidently follows that c_1 and c_2 in (18) should be complex conjugates; that is, we may write

$$c_1 = -(A + iB) \quad \text{and} \quad c_2 = -(A - iB),$$

where $A \geqq 0$. Thus the formula,

$$\varphi(u, 0) = \exp\left[-\left(A + iB\frac{u}{|u|}\right)|u|^{\lambda}\right]$$

holds for all u in the interval $|u| < \delta$.

It can be easily shown that $\delta = +\infty$, since from the continuity of the characteristic function, we have $\varphi(\pm\delta, 0) \neq 0$, which contradicts the assumption that δ is the smallest of the absolute value of the real root of $\varphi(u, 0)$. Hence, the characteristic function of the distribution of ξ is given by

$$\varphi(u, 0) = \exp\left[-\left(A + iB\frac{u}{|u|}\right)|u|^{\lambda}\right]. \tag{19}$$

Now, it should be noted that the case $A = 0$ should be excluded, since when $A = 0$, $|\varphi(u, 0)| = 1$ for all u; this leads to the trivial case that the whole mass of the distribution is concentrated at a single point.

It is already pointed out in (15) that λ does not involve a, so that on solving $\lambda = a[\theta'(a)/\theta(a)]$ as a differential equation in a, we get

$$\theta(a) = K|a|^{\lambda}. \tag{20}$$

Hence, we have from (13)

$$\varphi_1(au) = \exp\left[-K\left(A + iB\frac{u}{|u|}\right)|au|^{\lambda}\right], \tag{21}$$

where $K > 0$ for the same reason as $A > 0$.

Next, we shall show that $1 < \lambda \leqq 2$. If $\lambda \leqq 1$, the first derivatives of both $\varphi_1(u)$ and $\varphi(u, 0)$ fail to exist at the origin, which means that the first moments of ξ and x do not exist, contrary to the assumption of our theorem. On the other hand, if $\lambda > 2$, the second derivatives of both the functions $\varphi_1(u)$ and $\varphi(u, 0)$ exist and vanish at the origin. In this case, the second moments of both ξ and x exist and are equal to zero, which means that the whole mass of the distribution is concentrated at the point $\xi = x = 0$, and we have $\varphi_1(u) = \varphi(u, 0) = 1$ for all u (c.f. Cramer [1]). Now, from Lévy and Khintchine [6], it follows evidently that the characteristic functions (19) and (21) uniquely determine the distribution function of a stable law with finite expectation when and only when the parameters A, B, K, and λ satisfy the restrictions

$$\begin{cases} A > 0; \quad K > 0; \quad 1 < \lambda \leqq 2; \\ \left|B\cos\left(\frac{\pi}{2}\lambda\right)\right| \leqq A\sin\left(\frac{\pi}{2}\lambda\right). \end{cases} \tag{22}$$

PROOF OF SUFFICIENCY. We have to show that if the distribution functions of ξ and x are characterized by (19) and (21), respectively, with the parameters satisfying the restrictions as listed in (22), and further that if the first moment of η exists and the regression of η on ξ exists and is given by $E_\xi(\eta) = \beta_0\xi$, then the regression of $Y = cx + \eta$ on $X = ax + \xi$ should also be linear, where x is independent of the joint distribution of ξ and η.

Here we have

$$\ln \varphi_1 (au) = K|a|^\lambda \ln \varphi(u, 0). \tag{23}$$

Then we have

$$\frac{\varphi_1'(au)}{\varphi_1(au)} = \frac{1}{a} \cdot \frac{d \ln \varphi_1(au)}{du} = \frac{1}{a} K \mid a \mid^\lambda \frac{d \ln \varphi(u, 0)}{du},$$

so that we have

$$\varphi_1' (au)\, \varphi(u, 0) = \frac{1}{a} K|a|^\lambda \varphi_1(au)\, \varphi'(u, 0). \tag{24}$$

Then, if $\Phi(u, v)$ stands for the characteristic function of the joint cumulative distribution of X and Y, we have, on substituting the value of $\varphi_1'(au)\varphi(u, 0)$ as in (24) in (4) and (5) above,

$$\begin{cases} \left.\dfrac{\partial\Phi(u, v)}{\partial v}\right]_{v=0} = \left(\dfrac{c}{a} K\, |a|^\lambda + \beta_0\right) \varphi_1(au)\varphi'(u, 0), \\ \dfrac{d\Phi(u, 0)}{du} = (1 + K\, |a|^\lambda)\varphi_1(au)\varphi'(u, 0), \end{cases} \tag{25}$$

so that

$$\left.\frac{\partial\Phi(u, v)}{\partial v}\right]_{v=0} = \frac{\beta_0 + \dfrac{c}{a} K\, |a|^\lambda}{1 + K\, |a|^\lambda} \frac{d\Phi(u, 0)}{du}. \tag{26}$$

Then the proof follows at once from (26), using the lemma.

It is also interesting to note in this connection that if we further assume that either ξ or x has a finite variance, that is, that the second derivative of either (19) or (21) exists at the origin, then λ should be equal to 2, and hence both x and ξ should be normally distributed.

COROLLARY 1. (*The problem of Ragnar Frisch.*) *In the problem of Ragnar Frisch, which has been solved independently by Rao* [8], [9] *and by Fix* [3], *it has been assumed that x, ξ, and η are mutually independent chance variables. Thus, it may be treated as a special case of Theorem* 2, *above, by putting $\beta_0 = 0$.*

COROLLARY 2. (*Generalisation of Darmois' Theorem.*) *If x and y are two independent chance variables with finite expectations such that the regression of $Y = cx + dy$ ($d \neq 0$) on $X = ax + by$ ($b \neq 0$) exists and is linear for all a contained in a closed interval (a_1, a_2), where either $a_1 < a_2 < 0$ or $0 < a_1 < a_2$, and for some* *c for which the relationship $bc \neq ad$ is satisfied for all a in the interval, then both x and y should belong to the class of stable law with finite expectation.*

This may be treated as a special case of Theorem 2, above, by taking $\xi = by$ and $\eta = dy$. Finally, we shall construct a simple counter-example to show that the theorem is not true when the regression of Y on X is linear for some fixed a.

For this purpose, let us take

$$\ln \varphi(t) = \int_0^\infty (\cos |t|\, x - 1) \frac{e^{-\sin^2 \ln x}}{x^{1+\delta}}\, dx \qquad (1 < \delta < 2). \tag{27}$$

We shall show that $\varphi(t)$ in (27) represents the characteristic function of a symmetric infinitely divisible law with a finite first moment which is assumed to be zero. First of all, we note easily that $\varphi(t)$ in (27), being real, represents the characteristic function of a symmetric law. Now, following the notations given by Loève [7], we define

$$G(x) = \begin{cases} -\displaystyle\int_x^\infty \frac{e^{-\sin^2 \ln x}}{x^{1+\delta}}\, dx & \text{if } x > 0, \\ \displaystyle\int_{-\infty}^x \frac{e^{-\sin^2 \ln |x|}}{|x|^{1+\delta}}\, dx & \text{if } x < 0, \end{cases}$$

where $1 < \delta < 2$.

It can be easily verified that $G(x)$ satisfies all the conditions stated in Loève's representation formula for the infinitely divisible law. Hence, $\varphi(t)$ in (27), above, is the characteristic function of a symmetric infinitely divisible law.

Using the transformation $|t|x = u$, (27) reduces to

$$\ln \varphi(t) = |t|^\delta \int_0^\infty (\cos u - 1) \frac{\exp\{-\sin^2[\ln u - \ln |t|\,]\}}{u^{1+\delta}}\, du, \qquad 1 < \delta < 2. \tag{28}$$

Now the first derivative of $\varphi(t)$ in (28) exists at the origin, so that the first moment exists and is equal to zero.

Again, from (28), we have

$$\ln \varphi(at) = |at|^\delta \int_0^\infty (\cos u - 1) \frac{\exp\{-\sin^2[\ln u - \ln |t| - \ln |a|\,]\}}{u^{1+\delta}}\, du, \qquad 1 < \delta < 2. \tag{29}$$

Then, using (28) and (29) together, we have at the point $a = e^{2k\pi}$, where k takes any of the values $\pm 1, \pm 2, \pm 3 \cdots$,

$$\ln \varphi(at) = |a|^\delta \ln \varphi(t). \tag{30}$$

If the characteristic functions of the cumulative distributions of x and ξ are given by (27), above, then proceeding exactly in the same way as in (23), (24), (25), and (26), it can be shown that the regression of $Y = cx + \eta$ on $x = ax + \xi$

is linear for some fixed a, where $a = e^{2k\pi}$ and k is any one of the numbers ± 1, ± 2, ± 3, $\cdots$.

In conclusion, the author expresses his thanks to the Referee for some helpful comments.

REFERENCES

1. H. Cramér, *Mathematical methods of statistics*, Princeton University Press, 1946, p. 91.
2. G. Darmois, "Sur une propriete caracteristique de la loi de probabilite de Laplace," *C. R. Acad. Sci. Paris*, Vol. 232 (1951), pp. 1999–2000.
3. E. Fix, "Distributions which lead to linear regressions," *Proceedings of the First Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, 1949, pp. 79–91.
4. B. V. Gnedenko, "On a theorem of S. N. Bernstein," *Izvestiya Akad. Nauk SSSR. Ser. Mat.*, Vol. 12 (1948), pp. 97–100.
5. M. Kac, "A characterization of the normal distribution," *Amer. J. Math.*, Vol. 61 (1939), pp. 726–728.
6. P. Lévy and A. Khintchine, "Sur les loi stables," *C. R. Acad. Sci. Paris*, Vol. 202 (1936), pp. 374–376.
7. M. Loève, "Fundamental limit theorems of probability theory," *Ann. Math. Stat.*, Vol. 21 (1950), pp. 329.
8. C. R. Rao, "Note on a problem of Ragnar Frisch," *Econometrica*, Vol. 15 (1947), pp. 245–249.
9. C. R. Rao, "A correction to note on a problem of Ragnar Frisch," *Econometrica*, Vol. 17 (1949), p. 212.
10. Colette Rothschild and Edith Mourier, "Sur les lois de probabilité à regression lineaire et écart type lié constant," Comptes Rendus 225 (1947), pp. 245–249.