DISTRIBUTIONS WHICH LEAD TO LINEAR REGRESSIONS

EVELYN FIX

UNIVERSITY OF CALIFORNIA, BERKELEY

1. In 1936, during the Oxford Conference of the Econometric Society, Ragnar Frisch proposed the following question.

Suppose it is known that two random variables X and Y have the following composition:

$$\left.\begin{aligned} X &= a\xi + \alpha \\ Y &= b\xi + \beta \end{aligned}\right\} \tag{1}$$

where ξ, α, and β are some mutually independent random variables and a and b are certain constant coefficients, the values of which are unknown. What are the conditions under which the regression of Y on X, and also that of X on Y, is linear, irrespective of the values of a and b?

A partial answer to the question was given by H. V. Allen [**1**].[1] Miss Allen proved this theorem: provided the first two moments of β and all the moments of ξ and α exist, the necessary and sufficient condition for linearity of regression of Y on X, whatever may be a and b, is that both ξ and α should be normally distributed. The proof was based on the construction of an infinite sequence of polynomials, normal and orthogonal with respect to the elementary probability law of X, and therefore the existence of all moments as postulated was necessary in this manner of proof. However, as the author herself points out, the condition that all moments exist is restrictive and makes her answer only a partial one. It may be noted that the distributions considered in many practical problems do have finite moments of all orders.

The purpose of this article is to consider the same problem under assumptions regarding the moments which are less restrictive than those of Miss Allen.

2. First, a precise interpretation must be given the problem of Ragnar Frisch. Let X and Y be any two random variables and let $F_{X,Y}(x,y)$ and $F_X(x)$ stand for the joint cumulative distribution of both variables and for the marginal cumulative distribution of X, respectively. Thus, for all real x and y,

$$F_{X,Y}(x,y) = P\{(X \leqq x)(Y \leqq y)\} \tag{2}$$

and

$$F_X(x) = P\{X \leqq x\} = \lim_{y\to\infty} F_{X,Y}(x,y)\,. \tag{3}$$

Statements concerning the regression of Y on X will be interpreted to presuppose the existence, for all real x except perhaps for a set of probability zero

[1] Boldface numbers in brackets refer to references at the end of the paper (see p. 91).

(almost all x), of a function $F_{Y|x}(y)$, representing the probability distribution of Y relative to the assumption that $X = x$ (relative distribution of Y given x), such that for all real x and y

$$F_{X,Y}(x,y) = \int_{-\infty}^{x} F_{Y|x}(y) dF_X(x) \,. \tag{4}$$

Then the regression of Y on X is defined as

$$Y(x) = \int_{-\infty}^{\infty} y \, dF_{Y|x}(y) \,, \tag{5}$$

provided the integral on the right-hand side is absolutely convergent.

Thus, the statement that the regression of Y on X is represented by a polynomial of the nth degree implies the existence of $F_{Y|x}(y)$ for almost all real values of x, the absolute convergence of the integral in (5), and the existence of $n + 1$ real numbers c_k, $k = 0, 1, 2, \cdots, n$, such that for almost all x

$$Y(x) = \int_{-\infty}^{\infty} y dF_{Y|x}(y) = \sum_{k=0}^{n} c_k x^k \,. \tag{6}$$

3. Again, let X and Y be two random variables for which the regression of Y on X exists. Let $\varphi_{X,Y}(t,\tau)$, $\varphi_X(t)$ and $\varphi_{Y|x}(\tau)$ denote the characteristic functions of the distributions $F_{X,Y}(x,y)$, $F_X(x)$, and $F_{Y|x}(y)$ respectively. Then

$$\begin{aligned} \varphi_{X,Y}(t,\tau) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{itx+i\tau y} \, dF_{X,Y}(x,y) \\ &= \int_{-\infty}^{\infty} \left\{ e^{itx} \int_{-\infty}^{\infty} e^{i\tau y} \, dF_{Y|x}(y) \right\} dF_X(x) \\ &= \int_{-\infty}^{\infty} e^{itx} \varphi_{Y|x}(\tau) \, dF_X(x) \,. \end{aligned} \tag{7}$$

Also, if the regression of Y on X exists, that is to say, if the integral in (5) is absolutely convergent, then it is well known that the derivative of the characteristic function $\varphi_{Y|x}(\tau)$ exists and, at $\tau = 0$,

$$\left. \frac{d\varphi_{Y|x}(\tau)}{d\tau} \right|_{\tau=0} = i \, Y(x) \,. \tag{8}$$

Lemma 1. *If the regression of Y on X exists and, besides, if the first moment $\mu(Y)$ of Y also exists, so that the integral*

$$\mu(Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y \, dF_{X,Y}(x,y) \tag{9}$$

is absolutely convergent, then $\varphi_{X,Y}(t,\tau)$ *is differentiable with respect to* τ *and*

$$\frac{\partial \varphi_{X,Y}(t,\tau)}{\partial \tau} = \int_{-\infty}^{\infty} e^{itx} \frac{d\varphi_{Y|x}}{d\tau}\, dF_X(x)\,. \tag{10}$$

The proof of Lemma 1 is very simple and reduces to showing that the integral in (10) is convergent uniformly in τ. For this purpose, notice that

$$\left| e^{itx} \frac{d\varphi_{Y|x}}{d\tau} \right| = \left| \int_{-\infty}^{\infty} y\, e^{i\tau y}\, dF_{Y|x}(y) \right| \leqq \int_{-\infty}^{\infty} |y|\, dF_{Y|x}(y) \tag{11}$$

and that, therefore, whatever be A $< B$,

$$\left| \int_{A}^{B} e^{itx} \frac{d\varphi_{Y|x}}{d\tau}\, dF_X(x) \right| \leqq \int_{A}^{B} \int_{-\infty}^{\infty} |y|\, dF_{X,Y}(x,y)\,. \tag{12}$$

The proof is completed by reference to the absolute convergence of integral (9).

In particular, the substitution of $\tau = 0$ in (10) and the use of (8) gives

$$\left. \frac{\partial \varphi_{X,Y}(t,\tau)}{\partial \tau} \right|_{\tau=0} = i \int_{-\infty}^{\infty} e^{itx} Y(x)\, dF_X(x)\,. \tag{13}$$

Lemma 2. *If the first moment of* Y *and also the* n *first moments of* X *exist and if, furthermore, the regression of* Y *on* X *exists and is represented by a polynomial of the nth degree* $\sum_{k=0}^{n} c_k x^k$, *then*

$$\left. \frac{\partial \varphi_{X,Y}(t,\tau)}{\partial \tau} \right|_{\tau=0} = \sum_{k=0}^{n} c_k i^{1-k} \frac{d^k \varphi_X}{dt^k}\,. \tag{14}$$

The existence of the derivatives on the right-hand side of (14) is guaranteed by the hypothesis that the first n moments of X exist. This implies that the integrals

$$\mu_k(x) = \int_{-\infty}^{\infty} x^k dF_X(x) \tag{15}$$

are absolutely convergent for $k = 1, 2, \cdots, n$. Then, as is well known,

$$\frac{d^k \varphi_X}{dt^k} = i^k \int_{-\infty}^{\infty} x^k e^{itx} dF_x(x)\,. \tag{16}$$

The proof of Lemma 2 is completed by reference to formula (13).

Formulae (13) and (14) play an essential role in the following. In the present paper, formula (14) will be used assuming $n = 1$ only.

4. With reference to the random variables X and Y defined in (1), it is easy to see that the statement of the problem of Ragnar Frisch implies the existence of the first moments of all three variables ξ, α, and β. In fact, if $b = 0$, then the

regression of Y on X reduces to the first moment of β and the assumption that this regression exists for all real b implies the existence of the first moment of β. Again, if the regression of Y on X exists when $a = 0$ and $b = 1$, this would require the existence of the first moment of ξ. The further assumption that the regression of X on Y exists leads to the conclusion that the first moment of α must exist.

This seems to justify the consideration of the problem of Ragnar Frisch under the restriction that the first moments of ξ, α, and β exist in the sense that the integrals representing these moments are absolutely convergent. It is obvious that no loss of generality will be incurred by assuming that the first moments of ξ, α, and β are all equal to zero. Under these general restrictions, the solution of the problem of Ragnar Frisch is given by the following three theorems. In these three theorems the variables X, Y, ξ, α, and β will be those of formula (1).

Theorem 1. *For the regression of Y on X to be linear for all values of 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 $b \neq 0$, it is necessary and sufficient that*

either (i) *ξ or α reduces to a constant, $\xi = 0$ or $\alpha = 0$ or both*

or (ii) *that the characteristic functions of ξ and α have the form*

$$\varphi_\xi(t) = e^{-k^2\left(u+iv\frac{t}{|t|}\frac{a}{|a|}\right)|t|^\nu}, \tag{17}$$

$$\varphi_\alpha(t) = e^{-\left(u+iv\frac{t}{|t|}\right)|t|^\nu} \tag{18}$$

with

$$1 < \nu \leqq 2, \qquad u > 0 \qquad \text{and} \qquad k^2 > 0\,. \tag{19}$$

The alternative (i) is trivial, and therefore it will be assumed in the following that neither ξ nor α reduces to a constant.

5. *Proof of necessity.* Assume that neither ξ nor α reduces to a constant and that the regression of Y on X is linear for all $a \epsilon (a_1, a_2)$ and for some $b \neq 0$. This latter assumption combined with the postulated equality to zero of the first moments of ξ, α, and β implies that the first moments of X and Y exist and are equal to zero and that there exists a real number c such that, for almost all real x

$$Y(x) = cx\,. \tag{20}$$

Let $\varphi_\xi(t)$, $\varphi_\alpha(t)$, and $\varphi_\beta(t)$ denote the characteristic functions of ξ, α, and β respectively. Then

$$\varphi_X(t) = \varphi_\xi(at)\varphi_\alpha(t) \tag{21}$$

and

$$\varphi_{X,Y}(t,\tau) = \varphi_\xi(at + b\tau)\varphi_\alpha(t)\varphi_\beta(\tau)\,. \tag{22}$$

Now formula (14) will be applied to (21) and (22). In doing so it will be remembered that the equality to zero of the first moment of β implies that the derivative $\varphi_\beta{}'(0) = 0$. It will also be remembered that, whatever be the variable, its characteristic function taken at $t = 0$ is necessarily equal to unity. Using (21), (22), and (14), we have

$$b\varphi_\xi'(at)\varphi_\alpha(t) = c\frac{d}{dt}[\varphi_\xi(at)\varphi_\alpha(t)] \tag{23}$$

or

$$(b - ac)\varphi_\xi'(at)\varphi_\alpha(t) = c\varphi_\xi(at)\varphi_\alpha'(t) \tag{24}$$

for all real values of t and for all $a\epsilon(a_1,a_2)$.[2]

It is evident that c cannot equal zero. Suppose, in fact, that $c = 0$. Then $b - ac \neq 0$ and, for all values of t and $a\epsilon(a_1,a_2)$

$$\varphi_\xi'(at)\varphi_\alpha(t) = 0\,. \tag{25}$$

Since $\varphi_\alpha(0) = 1$ and since characteristic functions are continuous, there exists a number $\delta > 0$ such that, for $|t| < \delta$,

$$\varphi_\alpha(t) \neq 0\,. \tag{26}$$

It follows from (25) that for $|t| < \delta/a_0$

$$\varphi_\xi'(t) = 0\,, \tag{27}$$

where a_0 is the larger of the two numbers $|a_1|$ and $|a_2|$. This implies that the second derivative of $\varphi_\xi(t)$ at $t = 0$ must exist and be equal to zero. (See Cramér [2], p. 90.) From this it would follow that ξ is a constant equal to zero, contrary to the assumption. Similarly, $b - ac$ cannot be equal to zero without α reducing to a constant.

It may happen that $\varphi_\xi(t)$ and/or $\varphi_\alpha(t)$ have real roots. In that case let T be the smallest of the absolute values of the roots of either function. Since both $\varphi_\xi(t)$ and $\varphi_\alpha(t)$ are continuous and equal unity at $t = 0$, it follows that $T > 0$. Then for $|t| < T$ neither the function $\varphi_\xi(t)$ nor the function $\varphi_\alpha(t)$ is ever equal to zero. If neither $\varphi_\xi(t)$ nor $\varphi_\alpha(t)$ has any real roots, then T will stand for $+\infty$. Further, let T_0 be the smaller of the two numbers T and T/a_0. Then for all values of $a\epsilon(a_1,a_2)$ and for all $|t| < T_0$ neither $\varphi_\xi(at)$ nor $\varphi_\alpha(t)$ ever vanishes. If T_0 is a finite number, then either $\varphi_\alpha(\pm T_0) = 0$ or there will be at least one value $a'\epsilon(a_1,a_2)$ for which $\varphi_\xi(\pm a'T_0) = 0$ or both. Restricting ourselves to the interval $|t| < T_0$, we can write

$$(b - ac)\frac{\varphi_\xi'(at)}{\varphi_\xi(at)} = c\frac{\varphi_\alpha'(t)}{\varphi_\alpha(t)} \tag{28}$$

or

$$(b - ac)\frac{\partial \log \varphi_\xi(at)}{\partial t} = ac\frac{d \log \varphi_\alpha(t)}{dt}\,. \tag{29}$$

[2] This equation will be satisfied if $a = 1$ and ξ and α both have the same distribution function, regardless of its form, provided that the first moments exist as postulated. In this case the value of c will be $b/2$. (See J. F. Kenney [5].)

Integration of (29) with respect to t gives

$$(b - ac) \log \varphi_\xi(at) = ac \log \varphi_\alpha(t) + \log A \tag{30}$$

or

$$\varphi_\xi(at) = A[\varphi_\alpha(t)]^{\frac{ac}{b-ac}} \tag{31}$$

where $\log A$ is the constant of integration. But $A = 1$, since every characteristic function is equal to unity at $t = 0$. It will be remembered that (31) holds for all $a\epsilon(a_1,a_2)$ and for $|t| < T_0$.

Since $\varphi_\xi(at)$ is differentiable with respect to a, it follows that the exponent in the right hand side of (31), say

$$w = \frac{ac}{b - ac}, \tag{32}$$

must also be differentiable with respect to a. Denoting by w' the derivative of w, we have

$$t\varphi_\xi'(at) = w'\{\varphi_\alpha(t)\}^w \log \varphi_\alpha(t)\,. \tag{33}$$

Substitution in (33) of the value

$$\varphi_\xi'(at) = \frac{w}{a}\frac{\varphi_\alpha'(t)}{\varphi_\alpha(t)}\{\varphi_\alpha(t)\}^w \tag{34}$$

obtained from (28) and (31) gives

$$t\frac{\varphi_\alpha'(t)}{\varphi_\alpha(t)} = a\frac{w'}{w}\log \varphi_\alpha(t)\,. \tag{35}$$

Again, this equation is valid for $a\epsilon(a_1,a_2)$ and for $|t| < T_0$. There must exist an interval δ of values of t partial to $|t| < T_0$ and not including zero where the product $t \log \varphi_\alpha(t)$ never vanishes. In fact, since $\varphi_\alpha(t)$ cannot be equal to unity identically in any interval Δ including zero, however small, without α reducing to a constant, the interval $|t| < T_0$ must contain points t' where $\varphi_\alpha(t') \neq 1$. Since $\varphi_\alpha(t)$ is continuous, there must be a vicinity of t' where $\varphi_\alpha(t)$ is never equal to unity. Let δ denote the longest interval including t' and partial to $|t| \leqq T_0$ such that the product $t \log \varphi_\alpha(t)$ never vanishes within δ. However, then it will have to vanish at one of the boundaries, at least, of δ, which may be $t = 0$. Also the other boundary of δ may be $\pm T_0$.

Within the interval δ equation (35) may be divided by $t \log \varphi_\alpha(t)$ giving

$$\frac{\partial \log \log \varphi_\alpha(t)}{\partial t} = a\,\frac{w'}{w}\,\frac{d \log |t|}{dt}\,. \tag{36}$$

Integrating (36), we get the value of

$$\log \log \varphi_\alpha(t) = a\frac{w'}{w}\log |t| + \log K \tag{37}$$

valid for all the interval δ, which leads to

$$\varphi_a(t) = e^{K|t|^\nu} \tag{38}$$

with

$$\nu = a\frac{w'}{w} \tag{39}$$

and log K representing the constant of integration, possibly a complex number

$$K = -(u + iv)\,. \tag{40}$$

The constant K cannot reduce to zero without a being a constant. Also, since $|\varphi_a(t)| \leqq 1$, u must be greater than or equal to zero. Further on it will be shown that u is necessarily greater than zero. It is easy to see that neither ν nor K can depend on a. For this purpose it is sufficient to substitute in (37) two different values of t, say t_1 and t_2, both belonging to the interval δ. Then ν and K will appear as solutions of a system of two linear equations independent of a.

Since ν is an absolute constant and since (a_1,a_2) does not include zero, (39) implies

$$w = \frac{ac}{b - ac} = k^2|a|^\nu \tag{41}$$

with log k^2 standing for the constant of integration. Presently it will be shown that k^2 must be positive. Solving (41) for c we get

$$c = \frac{a}{|a|}\,\frac{bk^2|a|^{\nu-1}}{1 + k^2|a|^\nu}\,. \tag{42}$$

Using (31), (38), and (41), it follows that

$$\varphi_\xi(at) = e^{Kk^2|at|^\nu}\,. \tag{43}$$

It is seen from (43) that k^2 must be positive, or otherwise $|\varphi_\xi(at)|$ would have values greater than unity.

Formulae (38) and (43), valid within the whole interval δ not including zero and partial to $|t| < T_0$, imply that δ extends either from $-T_0$ to zero or from zero to $+T_0$. In fact the boundaries of δ are either zero and $\pm T_0$ or else such points $t' \neq 0$ where

$$\log \varphi_a(t') = K|t'|^\nu = 0\,. \tag{44}$$

Since $K \neq 0$, it follows that no such points exist. Thus formulae (38) and (43) hold either in $(-T_0,0)$ or in $(0,T_0)$. Since the values of characteristic functions taken at $\pm t \neq 0$ are necessarily conjugate, we may now determine both characteristic functions over the whole interval $|t| < T_0$. Suppose for example that $\delta = (0,T_0)$. Then within $(-T_0,0)$ we must have

$$\varphi_a(t) = \overline{\varphi_a(-t)} = e^{-(u-iv)|t|^\nu} \tag{45}$$

with the formula

$$\varphi_a(t) = e^{-\left(u+iv\frac{t}{|t|}\right)|t|^{\nu}} \tag{46}$$

being valid for the whole interval $|t| < T_0$. Similarly it is found that within $|t| < T_0$

$$\varphi_\xi(at) = e^{-k^2\left(u+iv\frac{t}{|t|}\right)|at|^{\nu}}. \tag{47}$$

Obviously, the functions $\varphi_a(t)$ and $\varphi_\xi(at)$ as in (46) and (47) satisfy the differential equation (29) throughout the interval $|t| < T_0$. The difference between the values of the functions in the negative and in the positive halves of this interval consists in the difference in the constants of integration, owing to the discontinuity at $t = 0$.

It is now possible to show that $T_0 = +\infty$. In fact T_0 was defined as the smallest positive number such that either $\pm T_0$ is a root of $\varphi_a(t)$ or such that $\pm a' T_0$ is a root of $\varphi_\xi(at)$. Since neither of these functions as determined by (46) and (47) has any roots at all, and since both of the functions must be continuous, it follows that $T_0 = +\infty$ and thus that formulae (46) and (47) are valid for $a \epsilon (a_1, a_2)$ and for all real values of t.

Substituting $at = \tau$ in (47), it is easily found that

$$\varphi_\xi(\tau) = e^{-k^2\left(u+iv\frac{\tau}{|\tau|}\frac{a}{|a|}\right)|\tau|^{\nu}} \tag{48}$$

Thus, if both a_1 and a_2 are positive

$$\varphi_\xi(t) = e^{-k^2\left(u+iv\frac{t}{|t|}\right)|t|^{\nu}}. \tag{49}$$

On the other hand, if $a_1 < a_2 < 0$, then

$$\varphi_\xi(t) = e^{-k^2\left(u-iv\frac{t}{|t|}\right)|t|^{\nu}}. \tag{50}$$

To conclude the proof of the necessity of the conditions stated in Theorem 1, it remains to show that $1 < \nu \leqq 2$ and that $u > 0$. The inequalities for ν follow from the fact that, if $\nu \leqq 1$, then the derivatives of $\varphi_a(t)$ and $\varphi_\xi(t)$ fail to exist at $t = 0$, which means that the first moments of a and ξ do not exist, contrary to the basic hypothesis. Also, should $\nu > 2$, then the second derivatives of $\varphi_a(t)$ and $\varphi_\xi(t)$ would exist and would vanish at $t = 0$. Should this be the case, then the second moments of a and ξ would exist and be equal to zero. But then a and ξ would have been constants equal to zero and their characteristic functions would have been $\varphi(t) \equiv 1$.

It remains to prove that u must be a positive number, not zero. This is achieved by noticing that should $u = 0$ then the modulus of the characteristic function of a

$$|\varphi_a(t)| \equiv \left|e^{iv\frac{t}{|t|}|t|^{\nu}}\right| \equiv 1 \tag{51}$$

for all real values of t, which is impossible. In fact, Cramér [**3**] has pointed out that if the modulus of a characteristic function of some variable z is equal to unity for some value $t_0 \neq 0$, then z must be a discrete variable whose distribution is a step function with discontinuities at all or at some points z_k of the form

$$z_k = 2k\pi/t_0 + z_0 \tag{52}$$

for $k = 0, \pm 1, \pm 2, \cdots$.

This is easy to prove as follows. Assume that the modulus of

$$\varphi_z(t) = \int_{-\infty}^{\infty} e^{itz} dF_z \tag{53}$$

is equal to unity for $t = t_0$. Then

$$\begin{aligned} |\varphi_z(t_0)|^2 &= \left| \int_{-\infty}^{\infty} (\cos t_0 z + i \sin t_0 z) dF_z \right|^2 \\ &= \left(\int_{-\infty}^{\infty} \cos t_0 z dF_z \right)^2 + \left(\int_{-\infty}^{\infty} \sin t_0 z dF_z \right)^2 \\ &= 1. \end{aligned} \tag{54}$$

Now the squares of the integrals above may be written as products and then as double integrals, e.g.:

$$\begin{aligned} \left(\int_{-\infty}^{\infty} \cos t_0 z dF_z(z) \right)^2 &= \int_{-\infty}^{\infty} \cos t_0 z dF_z(z) \cdot \int_{-\infty}^{\infty} \cos t_0 y dF_z(y) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cos t_0 z \cos t_0 y dF_z(z) dF_z(y). \end{aligned} \tag{55}$$

In this way equation (54) can be rewritten as

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cos t_0(z - y) dF_z(z) dF_z(y) = 1. \tag{56}$$

It is now obvious that should there be a set S of values of (y,z) such that

$$\iint_S dF_z(z) dF_z(y) > 0 \tag{57}$$

where $t_0(z - y) \neq 2k\pi$, then the integral in (56) would have a value less than unity. It follows therefore that the probability of the argument of the cosine in (56) being equal to a multiple of 2π is unity. Thus the possible values of the variable z must differ from each other by multiples of $2\pi/t_0$.

It follows that, should a characteristic function $\varphi_z(t)$ have its modulus equal to unity for two incommensurable values $t_1 \neq 0$ and $t_2 \neq 0$, then the possible values of z corresponding to the discontinuities of its distribution function $F_z(z)$ would have to belong to two sequences of numbers, say

$$z_0' \pm \frac{2k\pi}{t_1}, \qquad k = 0, 1, 2, \cdots \tag{58}$$

and

$$z_0'' \pm \frac{2k\pi}{t_2}, \qquad k = 0, 1, 2, \cdots . \tag{59}$$

But it is obvious that the sequences (58) and (59) may have no more than one element in common. Thus $|\varphi_z(t)| \equiv 1$ implies that z is a constant equal, say, to z_0 in which case

$$\varphi_z(t) = e^{itz_0} . \tag{60}$$

However, (60) is not a particular case of the general form of the characteristic function of α found above. Thus $u > 0$. This completes the proof of the necessity of conditions enumerated in Theorem 1.

6. *Proof of sufficiency.* To prove that under the conditions of Theorem 1 with the assumed existence of the first moment $\mu(\beta)$ of β, the regression of Y on X necessarily exists and is linear, it is sufficient to prove that the regression of ξ on X exists and is linear. In fact, it is evident that, if the regression of ξ on X is $\xi(x)$, then the regression of $Y = b\xi + \beta$ is

$$Y(x) = b\xi(x) + \mu(\beta) . \tag{61}$$

Therefore in further reasoning it will be assumed that

$$Y = \xi . \tag{62}$$

The proof is based essentially on formula (13). In order to use this formula it must first be established that, if the characteristic functions of ξ and α are given by formulae (17) and (18) with restrictions (19), then for almost all values of x the regression $Y(x)$ exists and that the first moment of $Y = \xi$ exists. All this is implied by the particular form of the characteristic functions $\varphi_\alpha(t)$ and $\varphi_\xi(t)$ with $u > 0$, $k^2 > 0$ and $1 < \nu \leqq 2$. It is noticed first that under these conditions the derivatives of both characteristic functions exist for all values of t, tend to zero as $t \to 0$ and as $t \to \infty$, these derivatives tend to zero faster than an arbitrary negative power of t. According to the recent result of Fortet [**4**], these conditions are sufficient for the existence of the first moments. Thus, if α and ξ have characteristic functions (46) and (48), they must possess first moments equal to zero.

Furthermore, the characteristic functions (46) and (48) are absolutely integrable from $-\infty$ to $+\infty$. This implies that the distribution functions $F_\xi(y)$ and $F_\alpha(z)$ possess continuous derivatives, say $p_Y(y)$ and $p_\alpha(z)$, representing the frequency functions of the two variables $Y = \xi$ and α. These frequency functions determine the joint frequency function of Y and $X = a\xi + \alpha$, namely

$$p_{X,Y}(x,y) = p_\alpha(x - ay)p_Y(y)\,. \tag{63}$$

The relative frequency function of Y given x, $p_{Y|x}(y)$, is then given by

$$p_{Y|x}(y) = \frac{p_\alpha(x - ay)p_Y(y)}{\int_{-\infty}^{\infty} p_\alpha(x - ay)p_Y(y)dy} \tag{64}$$

and the regression function $Y(x)$ is given by

$$Y(x) = \frac{\int_{-\infty}^{\infty} yp_\alpha(x - ay)p_Y(y)dy}{\int_{-\infty}^{\infty} p_\alpha(x - ay)p_Y(y)dy}\,. \tag{65}$$

It is evident that $Y(x)$ must exist for almost all values of x since otherwise the first moment of Y would fail to exist.

Thus the conditions of validity of formula (13) are established. Since the frequency function of X exists, say

$$p_X(x) = \int_{-\infty}^{\infty} p_\alpha(x - ay)p_Y(y)dy\,, \tag{66}$$

formula (13) may be written as

$$\left.\frac{\partial \varphi_{X,Y}(t, \tau)}{\partial \tau}\right|_{\tau=0} = i\int_{-\infty}^{\infty} e^{itx}Y(x)p_X(x)dx\,. \tag{67}$$

Using (46) and (49) or various previous formulae satisfied by the characteristic functions $\varphi_\alpha(t)$ and $\varphi_\xi(t)$, easy algebra gives

$$\left.\frac{\partial \varphi_{X,Y}(t, \tau)}{\partial \tau}\right|_{\tau=0} = \frac{k^2|a|^\nu}{a(1 + k^2|a|^\nu)}\,\frac{d\varphi_X(t)}{dt} \tag{68}$$

so that, owing to (67), the right-hand side of (68) appears to be a Fourier transform of the product

$$iY(x)p_X(x)\,. \tag{69}$$

Owing to the particular form of the characteristic functions $\varphi_\alpha(t)$ and $\varphi_\xi(t)$, and especially owing to the fact that $u > 0$ and $k^2 > 0$, the squares of the moduli of $\varphi_X(t)$ and of $d\varphi_X/dt$ are integrable from $-\infty$ to $+\infty$. Therefore (see Widder [6], p. 202) we may write

$$iY(x)p_X(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{k^2|a|^\nu}{a(1 + k^2|a|^\nu)}\,\frac{d\varphi_X}{dt}\,e^{-itx}dt \tag{70}$$

and

$$p_X(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \varphi_X(t)e^{-itx}dt\,. \tag{71}$$

Integrating (70) by parts and using (71), we have

$$iY(x)p_X(x) = \frac{1}{2\pi}\,\frac{k^2|a|^\nu}{a(1+k^2|a|^\nu)}\left\{\varphi_X(t)e^{-itx}\Big|_{-\infty}^{\infty} + ix\int_{-\infty}^{\infty}\varphi_X(t)e^{-itx}dt\right\}$$

$$= i\,\frac{k^2|a|^\nu}{a(1+k^2|a|^\nu)}\,xp_X(x) \tag{72}$$

or

$$Y(x) = \frac{k^2|a|^\nu}{a(1+k^2|a|^\nu)}\,x \tag{73}$$

for every x where $p_X(x) > 0$. This completes the proof of Theorem 1.

7. *Theorem* 2. *With the previous notation holding, a necessary and sufficient condition that $Y(x) = cx$ for all values of a within an interval (a_1,a_2) such that $a_1 < 0 < a_2$ is that*

$$\varphi_\xi(t) = e^{-k^2u|t|^\nu} \tag{74}$$

and

$$\varphi_\alpha(t) = e^{-u|t|^\nu} \tag{75}$$

with

$$1 < \nu \leqq 2, \qquad u > 0 \qquad \textit{and} \qquad k^2 > 0\,. \tag{76}$$

Theorem 2 follows easily from Theorem 1. The new assumption implies that formulae (49) and (50) must be equivalent. Hence $v = 0$.

8. *Theorem* 3. *If the variables ξ, α, and β satisfy the conditions of Theorem 1 and if, besides, the second moment of either ξ or α is known to exist, then both ξ and α must be normally distributed.*

Suppose, for example, that the second moment of α exists. Let it be $\sigma^2 > 0$. For every t,

$$\frac{d^2\log\varphi_\alpha(t)}{dt^2} = -\left(u + iv\frac{t}{|t|}\right)\nu(\nu-1)|t|^{\nu-2}\,. \tag{77}$$

The existence of the second moment implies that as $t \to 0$ this expression must tend to $-\sigma^2$. This is possible only when $\nu = 2$, $v = 0$, and $2u = \sigma^2$. Thus in this case

$$\varphi_\alpha(t) = e^{-\frac{\sigma^2t^2}{2}} \tag{78}$$

which is the characteristic function of the normal law. It follows easily from this that ξ also must be normally distributed. Similar reasoning applies when the second moment of ξ is known to exist.

In conclusion, I wish to express my indebtedness to Professor J. Neyman and to Professor G. Pólya for their valuable suggestions and criticisms.

REFERENCES

1. ALLEN, H. V. "A theorem concerning the linearity of regression," *Stat. Res. Mem.*, vol. 2 (1938), pp. 60–68.
2. CRAMÉR, HARALD. *Mathematical Methods of Statistics.* Princeton University Press (1946).
3. ———. *Random Variables and Probability Distributions.* Cambridge University Press. (1937). Cf. p. 26.
4. FORTET, ROBERT. "Calcul des moments d'une fonction de répartition a partir de sa caractéristique," *Bull. Sci. Math.*, t. 68 (1944), pp. 117–131.
5. KENNEY, J. F. "The regression systems of two sums having random elements in common," *Annals of Math. Stat.*, vol. 10 (1939), pp. 70–73.
6. WIDDER, DAVID VERNON. *The Laplace Transform.* Princeton University Press (1941).