UNIVERSITY OF CALIFORNIA PUBLICATIONS
IN STATISTICS, 1949–1955

Evelyn Fix, "Tables of Noncentral χ^2," Vol. 1, No. 2, pp. 15–19. Price, 25 cents
J. Neyman, "On the Problem of Estimating the Number of Schools of Fish," Vol. 1, No. 3, pp. 21–36. Price, 25 cents
Harry M. Hughes, "Estimation of the Variance of the Bivariate Normal Distribution," Vol. 1, No. 4, pp. 37–52. Price, 35 cents
Michel Loève, "On Sets of Probability Laws and Their Limit Elements," Vol. 1, No. 5, pp. 53–88. Price, 50 cents
E. W. Barankin and J. Gurland, "On Asymptotically Normal, Efficient Estimators: I," Vol. 1, No. 6, pp. 89–130. Price, 50 cents
Lucien LeCam, "On Some Asymptotic Properties of Maximum Likelihood Estimates and Related Bayes' Estimates," Vol. 1, No. 11, pp. 277–330. Price, 75 cents
Paul Lévy, "Random Functions: General Theory with Special Reference to Laplacian Random Functions," Vol. 1, No. 12, pp. 331–390. Price, 50 cents
Werner R. Leimbacher, "On Some Classes of Sequential Procedures for Obtaining Confidence Intervals of Given Length," Vol. 2, No. 1, pp. 1–22. Price, 35 cents
Emanuel Parzon, "On Uniform Convergence of Families of Sequences of Random Variables," Vol. 2, No. 2, pp. 23–53. Price, 50 cents
Jean Bass, "Space and Time Correlations in a Turbulent Fluid, Part I," Vol. 2, No. 3, pp. 55–84. Price, 50 cents
Jean Bass, "Space and Time Correlations in a Turbulent Fluid, Part II: Space and Time Spectrum," Vol. 2, No. 4, pp. 85–98. Price, 25 cents
Harold Cramér, "On Some Questions Connected with Mathematical Risk," Vol. 2, No. 5, pp. 99–124. Price, 35 cents
Charles Kraft, "Some Conditions for Consistency and Uniform Consistency of Statistical Procedures," Vol. 2, No. 6, pp. 125–142. Price, 25 cents

Address Orders to
UNIVERSITY OF CALIFORNIA PRESS
Berkeley 4, California

ON THE EXISTENCE OF LINEAR REGRESSION IN LINEAR STRUCTURAL RELATIONS

BY
THOMAS FERGUSON

UNIVERSITY OF CALIFORNIA PUBLICATIONS IN STATISTICS
Volume 2, No. 7, pp. 143–166

UNIVERSITY OF CALIFORNIA PRESS
BERKELEY AND LOS ANGELES
1955

UNIVERSITY OF CALIFORNIA PUBLICATIONS IN STATISTICS
EDITORS: JERZY NEYMAN, L. M. LECAM, HENRY SCHEFFÉ

Volume 2, No. 7, pp. 143–166

Submitted by editors May 16, 1955

Issued December 30, 1955

Price, 35 cents

UNIVERSITY OF CALIFORNIA PRESS
BERKELEY AND LOS ANGELES
CALIFORNIA

CAMBRIDGE UNIVERSITY PRESS
LONDON, ENGLAND

PRINTED IN THE UNITED STATES OF AMERICA

ON THE EXISTENCE OF LINEAR REGRESSION IN LINEAR STRUCTURAL RELATIONS

BY

THOMAS FERGUSON

1. Introduction. This paper is concerned with the linearity of the multiple regression of one variable on several others when the variables are connected by a linear structural relation. Extensions are made in several directions of certain results of other authors, namely, H. V. Allen [**1**], E. Fix [**4**], C. R. Rao [**16**], and D. V. Lindley [**13**]. Also a relation to identifiability is considered.

The problem of existence of a linear regression in a linear structural relation was first proposed in 1936 by Ragnar Frisch in a form which may be stated as follows. Let X_0 and X_1 be the observable random variables of a linear structural relation,

$$X_0 = a_0\xi + \eta_0 \qquad (1.1)$$
$$X_1 = a_1\xi + \eta_1$$

where ξ, η_0 and η_1 are independent random variables, and a_0 and a_1 are unknown constants. What are necessary and sufficient conditions on the distributions of the variables ξ, η_0, and η_1 in order that the regression of X_0 on X_1 and also that of X_1 on X_0 be linear for all values of the constants a_0 and a_1?

Miss Allen found, under the assumption of the existence of all moments of ξ and η_1, that a necessary and sufficient condition for the regression of X_0 on X_1 to be linear whatever be the value of a_1 is that both ξ and η_1 be normal random variables.

Miss Fix later completely solved the problem of Frisch. Her result, which contains that of Miss Allen, avoids the assumption of the existence of any moments other than the first, whose existence is implicit in Frisch's formulation of the problem, and requires the regression to be linear only for values of a_1 in an interval however small. For the sake of completeness we shall state this result in its entirety. We use $\phi_Y(t)$ to denote the characteristic function of a random variable Y; that is, $\phi_Y(t) = \int_{-\infty}^{\infty} e^{ity}\, dP[Y \leqq y]$. Without loss of generality we assume that $E\xi = E\eta_1 = 0$.

MISS FIX'S THEOREM. *In order that the regression of* X_0 *with* $a_0 \neq 0$ *on* X_1 *be a linear function of* X_1 *whatever be the value of* a_1 *in some interval* (c_1, c_2), *it is necessary and sufficient that*

$$\phi_\xi(t) = \exp\left[-\left(u + iv\frac{|t|}{t}\right)|t|^\nu\right]$$
$$\phi_{\eta_1}(t) = \exp\left[-k\left(u + iv\frac{|t|}{t}\frac{|a_1|}{a_1}\right)|t|^\nu\right]$$

where $1 < \nu \leqq 2$, $u > 0$, $k > 0$ *and* $\left|v \cos \frac{\pi\nu}{2}\right| \leqq u \sin \frac{\pi\nu}{2}$ *and where* $v = 0$ *should the interval* (c_1, c_2) *contain the origin.*

This paper was prepared with the support of the Office of Ordnance Research, United States Army, under Contract DA-04-200-ORD-171.

Random variables with a characteristic function of the above form are called stable, and symmetric stable if they have no imaginary part (that is, if $v = 0$).

If the second moment of either ξ or η_1 is assumed to exist so that the second derivative of the characteristic function exists at the origin, one easily sees that ν must equal 2; hence one obtains the easy corollary that in this case both ξ and η_1 must be normal variables. Thus Miss Fix's theorem implies the result of Miss Allen. The extension of this theorem and its corollary, requiring the regression to be linear for only two values of a_1, will be found in section 3.

In section 4 is found an extension in the dimensionality of the linear structural relation. It is assumed that there are $n + 1$ random variables, $X_0, X_1, \cdots, X_n$, each of which depends linearly on s random variables $\xi_1, \cdots, \xi_s$ and an independent error term. Necessary and sufficient conditions are found in order that the multiple regression of X_0 on $X_1, \cdots, X_n$ be linear irrespective of the values of the constants which appear in the linear forms. Other interesting remarks in connection with this problem may be found in the above-mentioned paper of Lindley. Using the above structure with $s = 1$ and $n > 1$, section 5 contains a theorem to the effect that if the regression is linear for just one set of values of the constants, then all the variables concerned must be normal. Next follows a theorem on linear regression when the error terms follow a joint stable distribution.

Finally, implications of these theorems on identifiability in a linear structural relation are considered, and two theorems on identifiability are proved. There is an indication given that linearity of regression and identifiability of the slope parameter in a linear structural relation are contradictory assumptions. Thus doubt is cast on the validity of certain methods used in factor analysis.

2. Preliminary lemmas. Let $X_0, X_1, \cdots, X_n$ be the observable random variables and $\xi_1, \cdots, \xi_s$ the latent random variables of a linear structural relation. Explicitly, let

$$X_0 = a_{01}\xi_1 + \cdots + a_{0s}\xi_s + \eta_0$$

$$X_1 = a_{11}\xi_1 + \cdots + a_{1s}\xi_s + \eta_1 \tag{2.1}$$

$$\cdots$$

$$X_n = a_{n1}\xi_1 + \cdots + a_{ns}\xi_s + \eta_n$$

where $\eta_0, \eta_1, \cdots, \eta_n$ are random variables, usually called the error terms, and the a_{jk} are unknown constants. Upon denoting the vectors $(X_0, X_1, \cdots, X_n)$, $(\xi_1, \cdots, \xi_s)$ and $(\eta_0, \eta_1, \cdots, \eta_n)$ by X', ξ', and η' respectively, and using A to denote the $(n + 1) \times s$ matrix of constants, equations (2.1) may be written simply as

$$X = A\xi + \eta .$$

Restating the problem, we are to find necessary and sufficient conditions on the distributions of ξ and η in order that for a certain number of values of the matrix A, there exist numbers $b_1, \cdots, b_n$, and c depending on A such that

$$E(X_0 | X_1, \cdots, X_n) = b_1X_1 + \cdots + b_nX_n + c . \tag{2.2}$$

For this and the following three sections, the four assumptions below on the distributions of ξ and η will be made tacitly.

Assumption 1. To conform to the practical case found in factor analysis applications in psychology and economics, it is assumed that the vector ξ is completely independent of the vector η.

Assumption 2. It is assumed that the first moments of both ξ and η exist. Without loss of generality we then assume $E(\xi) = E(\eta) = 0$. Thus c in equation (2.2) must be equal to zero.

The role of this assumption is to insure the existence of the regression of X_0 on $X_1, \cdots, X_n$. It will also permit us to take one derivative of the characteristic functions of ξ and η.

Assumption 3. It is assumed that the components of the vector η are completely independent among themselves.

Assumption 4. It is assumed that each ξ_k, $k = 1, \cdots, s$ and at least one of the η_j, $j = 1, \cdots, n$ are nondegenerate. We will take η_1 to be nondegenerate.

In section 6, we will treat the case $n = 1$, $s = 1$, where the components of the error term are not necessarily independent, but are known to follow a joint stable distribution.

The proofs of the theorems in the next sections use as a starting point a very simple necessary and sufficient condition for linear regression, involving the use of a differential equation connecting the characteristic functions of the variables concerned. This is contained in the following lemma which corresponds to lemma 2 in Miss Fix's paper.

Throughout we shall use $\phi_{Y_1,\ldots,Y_m}(t_1, \cdots, t_m)$ to denote the joint characteristic function of m random variables $Y_1, \cdots, Y_m$ and $\psi_{Y_1,\ldots,Y_m}(t_1, \cdots, t_m)$ to denote $\log \phi_{Y_1,\ldots,Y_m}(t_1 \cdots, t_m)$. Also $\psi_{Y_1,\ldots,Y_k',\ldots,Y_m}(t_1, \cdots, t_m)$ will be used to represent $\partial/\partial t_k[\psi_{Y_1,\ldots,Y_m}(t_1, \cdots, t_m)]$.

Lemma 1. *The following three statements are equivalent.*

(1) $E(X_0|X_1, \cdots, X_n) = b_1X_1 + b_2X_2 + \cdots + b_nX_n$ a.e.

(2) $\dfrac{\partial}{\partial t_0}\phi_{X_0,X_1,\ldots,X_n}(t_0, t_1, \cdots, t_n)|_{t_0=0} = \sum_{h=1}^{n} b_h \dfrac{\partial}{\partial t_h}\phi_{X_1,\ldots,X_n}(t_1, \cdots, t_n)$

(3) $$\sum_{k=1}^{s} \psi_{\xi_1,\ldots,\xi_k',\ldots,\xi_s}\left(\sum_1^n t_ja_{j1}, \cdots, \sum_1^n t_ja_{js}\right)\left(a_{0k} - \sum_{h=1}^{n} b_ha_{hk}\right) = \sum_{h=1}^{n} b_h\psi_{\eta_h}'(t_h) .$$

It should be noted that (1) and (2) are equivalent irrespective of the relation between the variables $X_0, X_1, \cdots, X_n$ provided that all first moments exist. Statement (3) is made in the sense that if both sides exist, they are equal. Of course, both sides do exist in a neighborhood of the origin where all the characteristic functions involved are not zero.

Proof. That statement (2) implies (3) is proved by rewriting the structural relation (2.1) in terms of characteristic functions and applying laborious computation and assumptions 1, 2, and 3. The reverse implication follows by retracing the steps involved and noting that whenever one side of (3) does not exist, then both sides of (2) are zero.

To see that (1) implies (2) we compute the left-hand side of (2).

$$\frac{\partial}{\partial t_0}\phi_{X_0,X_1,\cdots,X_n}(t_0, \cdots, t_n)\Big|_{t_0=0} = \frac{\partial}{\partial t_0} E\left[\exp i \sum_0^n t_j X_j\right]\Big|_{t_0=0}$$

$$= E\left[iX_0 \cdot \exp i \sum_1^n t_j X_j\right] = E\left[iE(X_0|X_1, \cdots, X_n) \cdot \exp i \sum_1^n t_j X_j\right].$$

The right-hand side of (2) is simply

$$\sum_{h=1}^n b_h E\left[iX_h \cdot \exp i \sum_1^n t_j X_j\right] = E\left[i\left(\sum_1^n b_h X_h\right) \cdot \exp i \sum_1^n t_j X_j\right].$$

Hence (1) obviously implies (2).

The reverse implication follows by rewriting (2)

$$E\left[iE(X_0|X_1, \cdots, X_n) \cdot \exp i \sum_1^n t_j X_j\right] = E\left[i\left(\sum_1^n b_h X_h\right) \cdot \exp i \sum_1^n t_j X_j\right]$$

or equivalently,

$$E\left[\left\{E(X_0|X_1, \cdots, X_n) - \sum_1^n b_h X_h\right\} \cdot \exp i \sum_1^n t_j X_j\right] = 0.$$

Hence from the unicity of the Fourier-Stieltjes transform in n-dimensions we have

$$E(X_0|X_1, \cdots, X_n) - \sum_1^n b_h X_h = 0 \qquad \text{a.e.}$$

It is well known that a sufficient condition for the regression of X_0 on $X_1, \cdots, X_n$ to be linear is that $X_0, X_1, \cdots, X_n$ be multivariate normal. A proof using densities is found in Cramér [**2**, p. 314]. An alternative method of proof which will work even if the distribution does not admit a density (the so-called singular normal) may be found using lemma 1. Hence we see that a sufficient condition for the regression of X_0 on $X_1, \cdots, X_n$ to be linear is that both ξ and η be multivariate normal. This fact will be used without explicit mention.

Certain properties of the stable distributions and in particular the multivariate stable distributions will be used frequently in the sequel. It will be convenient here to use a form of the stable law found in Gnedenko and Kolmogorov [**6**, p. 164]. The log of the characteristic function of every stable law admits the representation

$$\psi(t) = i\gamma t - c|t|^\nu\left\{1 + i\beta\frac{t}{|t|}\omega(t, \nu)\right\}$$

where ν, β, γ, and c are real constants, $0 < \nu \leqq 2$, $-1 \leqq \beta \leqq 1$, γ arbitrary, $c > 0$, and where

$$\omega(t, \nu) = \begin{cases} \tan\frac{\pi\nu}{2} & \text{if} \quad \nu \neq 1 \\ \frac{2}{\pi}\log|t| & \text{if} \quad \nu = 1. \end{cases}$$

The parameter ν is called the characteristic exponent. For a stable law with characteristic exponent ν, the absolute moments of all orders less than ν are finite, and if $\nu \neq 2$, the absolute moments of all orders greater than ν are infinite. The stable law with $\nu = 2$ is easily seen to be the normal. When the parameter β is zero, the stable law becomes symmetric about γ.

A well-known necessary and sufficient condition for random variables $Y_1, \cdots, Y_n$ to have a multivariate normal distribution is that every linear combination of them be normal. (See the note of M. Loéve in Lévy [**12**, p. 337].) It is most easily proved using characteristic functions. It may also be shown that a necessary and sufficient condition for $Y_1, \cdots, Y_n$ to be multivariate stable in the sense of Lévy [**11**, p. 221] is that every linear combination of the variables be stable. We shall use this as a definition of a multivariate stable law, and in particular, a symmetric multivariate stable law is defined to be one for which every linear combination of the variables is symmetric stable.

The following lemma gives a form of multivariate stable laws which will be used later in the paper. It should be noted that not every function of this form will be the log of a characteristic function. An exact representation for multivariate stable laws using an integral form may be found in Lévy [**11**].

LEMMA 2. *Random variables $Y_1, \cdots, Y_n$ have a multivariate stable distribution if and only if the logarithm of their joint characteristic function has the form*

$$\psi_{Y_1,\cdots,Y_n}(t_1, \cdots, t_n)$$
$$= i\gamma(t_1, \cdots, t_n) - c(t_1, \cdots, t_n)\{1 + i\beta(t_1, \cdots, t_n)\omega(1, \nu)\}$$

where

(i) $c(t_1u, \cdots, t_nu) = |u|^\nu c(t_1, \cdots, t_n) \qquad c > 0$

(ii) $\beta(t_1u, \cdots, t_nu) = \frac{u}{|u|}\beta(t_1, \cdots, t_n) \qquad -1 \leqq \beta \leqq 1$

(iii) $\gamma(t_1u, \cdots, t_nu) = \gamma(t_1, \cdots, t_n)u$
$$- c(t_1, \cdots, t_n)\beta(t_1, \cdots, t_n)|u|^\nu\frac{u}{|u|}[\omega(u, \nu) - \omega(1, \nu)].$$

This distribution is symmetric multivariate stable if and only if $\beta(t_1, \cdots, t_n) \equiv 0$.

Proof. To simplify the writing we use vector notation. Y' will denote the vector of variables $(Y_1, \cdots, Y_n)$, a' a vector of constants, and t an arbitrary constant.

Suppose the condition is satisfied. Then

$$\psi_{a'Y}(t) = \psi_Y(at) = i\gamma(at) - c(at)\{1 + i\beta(at)\omega(1, \nu)\}$$
$$= i\gamma(a)t - c(a)|t|^\nu\left\{1 + i\frac{t}{|t|}\beta(a)\omega(t, \nu)\right\}$$

which for every vector a is the log of the characteristic function of a stable distribution. If $\beta \equiv 0$, then $a'Y$ must be symmetric stable.

On the other hand, suppose Y is multivariate stable. Then for every vector a,

$$\psi_{a'Y}(t) = i\gamma(a)t - c(a)|t|^{\nu(a)}\left\{1 + i\beta(a)\frac{t}{|t|}\omega(t, \nu(a))\right\}. \tag{2.3}$$

Putting $t = 1$, we see that

$$\psi_Y(a) = i\gamma(a) - c(a)\{1 + i\beta(a)\omega(1, \nu(a))\}. \tag{2.4}$$

To complete the proof we will show that $\nu(a)$ is a constant on the set $E = \{a : c(a) \neq 0\}$, and that the functions c, β, and γ have the properties claimed.

Let Y_0 be an n-dimensional random vector with the same distribution as Y but independent of Y. Then $Y - Y_0$ is multivariate stable and

$$\psi_{Y-Y_0}(at) = -2c(a)|t|^{\nu(a)} = -2c(at). \tag{2.5}$$

Thus $c(a)$ is a continuous function of a, and $\nu(a)$ must be a continuous function of a on the open set E. It will be shown that $\nu(a)$ is constant on a dense set of the n-dimensional space; hence $\nu(a)$ would be constant on E. In fact, we will show there are at most $n - 1$ independent vectors b for which $\nu(b) > \min_a \nu(a)$. Suppose there are n independent vectors $b_1, \cdots, b_n$ and a vector a for which $\nu(a) < \min_j \nu(b_j)$. Let $B = (b_1, \cdots, b_n)$ be the nonsingular matrix composed of the b_j. There exists a vector d such that $Bd = a$; hence $a'Y = d'B'Y$.

$$a'Y = d_1(b_1'Y) + d_2(b_2'Y) + \cdots + d_n(b_n'Y).$$

An application of Minkowski's inequality will then give $\nu(a) \geq \min_j \nu(b_j)$ which is a contradiction. Hence, $\nu(a)$ is constant.

Equation (2.5) immediately gives (i). We may compute equation (2.4) at the point at and equate to (2.3). Equating the imaginary parts of the result gives

$$\gamma(at) = \gamma(a)t - c(a)|t|^{\nu}\left\{\beta(a)\frac{t}{|t|}\omega(t, \nu) - \beta(at)\omega(1, \nu)\right\}.$$

Hence (iii) will follow when we show (ii). We may expand $\gamma(at_1t_2)$ in two ways, first considering t_1t_2 as the constant multiplier and second considering t_1 and t_2 separately and applying the above formula each time. Equating the results gives

$$\beta(a)\frac{t_1}{|t_1|}\omega(t_1, \nu) - \beta(at_1)\omega(1, \nu)$$
$$= \frac{|t_2|^{\nu}}{t_2}\left\{\beta(a)\frac{t_1t_2}{|t_1t_2|}\omega(t_1t_2, \nu) - \beta(at_1)\frac{t_2}{|t_2|}\omega(t_1, \nu)\right\}.$$

By considering the cases $\nu = 1$ and $\nu \neq 1$ separately, (ii) follows easily. Finally, if $a'Y$ is symmetric stable, then $\beta(a) \equiv 0$.

There are several remarks to be made about lemma 2. First, if $\nu \neq 1$ then $\omega(u, \nu) = \omega(1, \nu)$ so that condition (iii) becomes simply $\gamma(t_1u, \cdots, t_nu) = \gamma(t_1, \cdots, t_n)u$. It may be shown further by centering the variables Y_j and noting that linear combinations of centered variables must be centered, that (iii) may be replaced by

$$\text{(iii}') \qquad \gamma(t_1, \cdots, t_n) = \sum_{i=1}^{n} \gamma_j t_j$$

for some arbitrary real constants γ_j. The same is true for symmetric variables whether $\nu = 1$ or not. Secondly, if $\nu = 1$, $\omega(1, \nu) = 0$ so that the characteristic function may be written in an obviously simpler form.

Functions which satisfy equation (i) of the lemma are called positive homogeneous of degree ν. It should be noted that by lemma 2, variables $Y_1, \cdots, Y_n$ have a multivariate normal distribution if and only if the log of their characteristic function is of the form

$$\psi_{Y_1,\cdots,Y_n}(t_1, \cdots, t_n) = i\sum_{i=1}^{n} \gamma_j t_j - c(t_1, \cdots, t_n),$$

where $c(t_1, \cdots, t_n)$ is a homogeneous function of degree 2.

3. Reduction in the number of constants for which the regression is assumed linear. In this section we treat a modification of the problem of Ragnar Frisch, requiring the regression to be linear for only two values of the constant a_1, and finding as a result that both ξ and η_1 must be semistable. It is assumed that the linear structure under consideration has the special form

$$\begin{aligned} X_0 &= a_0\xi + \eta_0 \\ X_1 &= a_1\xi + \eta_1 \end{aligned} \tag{1.1}$$

and that assumptions 1 to 4 are satisfied. To avoid cases where the regression is trivially linear we assume in the theorem and corollaries to follow that none of the constants a_0, a_1, primed or not, are zero.

THEOREM 1. *In order that the regression of X_0 on X_1 be linear for two pairs of values of the constants (a_0, a_1) and (a_0', a_1'), with $|a_1| \neq |a_1'|$, it is necessary and sufficient that either*

(i) *both ξ and η_1 be normal variables or*

(ii) $$\psi_\xi(t) = \int_{-\infty}^{\infty} (e^{itx} - 1 - itx)dN(x)$$
$$\psi_{\eta_1}(t) = \int_{-\infty}^{\infty} (e^{itx} - 1 - itx)dM(x)$$

where

$$M(x) = K\frac{|a_1|}{a_1}N\left(\frac{x}{a_1}\right), \qquad a_1 \neq 0, \quad K > 0$$

and

$$N(x) = \begin{cases} x^{-\nu}Q_1(\log x), & \textit{for} \quad x > 0 \\ |x|^{-\nu}Q_2(\log|x|), & \textit{for} \quad x < 0 \end{cases}$$

where Q_1 and Q_2 are periodic functions such that $N(x)$ (hence $M(x)$) is nondecreasing in both $(-\infty, 0)$ and $(0, \infty)$, and $1 < \nu < 2$. In the case $q = a_1/a_1' > 0$, Q_1 and Q_2 are both of period $\log q$; in the case $q < 0$, Q_1 and Q_2 are of period $2 \log |q|$ and $Q_1(y + \log |q|) = Q_2(y)$.

Proof. In the case $n = 1$, $s = 1$, condition (3) of lemma 1 becomes

$$\psi_\xi'(ta_1)(a_0 - ba_1) = b\psi_{\eta_1}'(t) .$$

Neither b nor $(a_0 - ba_1)$ may be zero without contradicting nondegeneracy. By integrating we have

$$\psi_\xi(ta_1)\frac{(a_0 - ba_1)}{ba_1} = \psi_{\eta_1}(t) . \tag{3.1}$$

Using the fact that the absolute value of a characteristic function is bounded by one we see that the constant $K = (a_0 - ba_1)/ba_1$ must be positive. This equation may be deduced once again using the primed structure, and upon equating the two expressions and letting $K' = (a_0' - b'a_1')/b'a_1' > 0$ we find

$$\psi_\xi\left(t\frac{a_1}{a_1'}\right) = \frac{K'}{K}\psi_\xi(t) , \tag{3.2}$$

or alternatively,

$$\psi_\xi(tq) = |q|^\nu\psi_\xi(t) . \tag{3.3}$$

This equation holds at least in a neighborhood of the origin, but since $q \neq 1$ it obviously holds everywhere on the real line.

Laws with characteristic functions whose logs satisfy this equation are called semistable by Lévy [**11**, p. 204]. Characteristic functions of such laws for q positive and unequal to one are shown to be representable in the form

$$\psi_\xi(t) = \left[-P_0(\log|t|) + i\frac{|t|}{t}P_1(\log|t|)\right]|t|^\nu , \tag{3.4}$$

where $0 < \nu \leqq 2$ and P_0 and P_1 are continuous real functions periodic of period $\log q$. The case $P_0 = P_1 = 0$ is ruled out by nondegeneracy; so that in order for the derivative of $\psi_\xi(t)$ to exist at the origin it is necessary that $\nu > 1$. Lévy uses the fact that ξ must be infinitely divisible to derive the exact form of the semistable law given in the theorem for ξ when q is positive. Using the unicity of the representation of an infinitely divisible law (see [**6**, p. 80]), the formula for M follows easily from equation (3.1).

If q is negative, we note that

$$\psi_\xi(tq^2) = |q|^\nu\psi_\xi(tq) = (q^2)^\nu\psi_\xi(t)$$

so that we get the same representation as before, but now Q_1 and Q_2 are periodic functions of period $2 \log |q|$. The final condition follows easily from equation (3.1).

In order to prove sufficiency, one merely checks that equations (3.1) and (3.2) are satisfied.

If the random variable ξ (similarly η_1) has a finite second moment, the second derivative of $\psi_\xi(t)$ must exist at the origin. Equation (3.4) will then give us that $\nu \geqq 2$. We have proved the following corollary.

COROLLARY 1. *In order that the regression of X_0 on X_1 be linear for two pairs of values of the constants (a_0, a_1) and (a_0', a_1') with $|a_1| \neq |a_1'|$, and that the second moment of either ξ or η_1 exist, it is necessary and sufficient that both ξ and η_1 be normal variables.*

To complete the picture we shall prove a second corollary, assuming the regression to be linear for three pairs of values of the constants satisfying a certain incommensurability relation. It is easily seen that this result contains Miss Fix's theorem.

COROLLARY 2. *In order that the regression of X_0 on X_1 be linear for three pairs of values of the constants (a_0, a_1) (a_0', a_1') and (a_0'', a_1'') such that*

$$\frac{\log|a_1| - \log|a_1'|}{\log|a_1'| - \log|a_1''|}$$

be an irrational number, it is necessary and sufficient that

$$\psi_\xi(t) = -\left(u + iv\frac{|t|}{t}\right)|t|^\nu$$

$$\psi_{\eta_1}(t) = -K\left(u + iv\frac{|t|}{t}\frac{|a_1|}{a_1}\right)|t|^\nu ,$$

where $1 < \nu \leqq 2$, $u > 0$, $K > 0$ and $|v \cos(\pi\nu/2)| \leqq u \sin(\pi\nu/2)$ and where $v = 0$ unless a_1, a_1' and a_1'' are all of the same sign.

Proof. Theorem 1 is applied twice to the form of semistable laws found in equation (3.4). We see that P_0 and P_1 are continuous and periodic both of periods $2 \log |a_1/a_1'|$ and $2 \log |a_1'/a_1''|$. Since these periods are incommensurable, it follows that P_0 and P_1 are constants. Hence the formula given for $\psi_\xi(t)$. The formula given for $\psi_{\eta_1}(t)$ follows easily from equation (3.1). The conditions $u > 0$, $K > 0$, and $|v \cos(\pi\nu/2)| \leqq u \sin(\pi\nu/2)$ are necessary in order that $\psi_\xi(t)$ and $\psi_{\eta_1}(t)$ represent logs of characteristic functions. (See Lévy [**11**, pp. 94–96].) The sufficiency follows of course from Miss Fix's theorem.

With regard to the preceding corollary, the author is indebted to an oral communication of Dr. LeCam who, however, supported the validity of the conclusion by arguments different from those presented here.

For the regression of X_0 on X_1 to be linear for only one value of (a_0, a_1), J. F. Kenney [**8**] has found that it is sufficient that $a_1\xi$ and η_1 have the same distribution regardless of its form. This may be seen from equation (3.1) with $b = a_0/2a_1$. This indicates the difficulty of finding necessary and sufficient conditions in this case.

4. Extension in the dimensionality of the structure. Consider now the general linear structural relation found in section 2, equations (2.1). It is proposed to find necessary and sufficient conditions for the regression of X_0 on $X_1, \cdots, X_n$ to be linear irrespective of the values of the constants in the matrix A. Corollary 2 of section 3 is used to obtain a necessary condition which, as will be seen, is almost good enough for sufficiency.

THEOREM 2. *In order that the regression of X_0 on $X_1, \cdots, X_n$ be a linear function of $X_1, \cdots, X_n$ irrespective of the values of the constants a_{jk}, it is necessary that the logarithms of the characteristic functions of $\xi_1, \cdots, \xi_s$ and η_1 through η_n be representable in the form*

$$\psi_{\xi_1,\cdots,\xi_s}(t_1, \cdots, t_s) = -g(t_1, \cdots, t_s)$$

$$\psi_{\eta_j}(t) = -K_j|t|^\nu, \qquad j = 1, \cdots, n,$$

where $g(t_1, \cdots, t_s)$ is a positive homogeneous real function of degree ν, $1 < \nu \leqq 2$, $K_j \geqq 0$ for $j = 2, \cdots, n$ and $K_1 > 0$.

Proof. We may choose a_{0k}, $k = 1, \cdots, s$ in equations (2.1) to make $a_{01}\xi_1 + \cdots + a_{0s}\xi_s$ nondegenerate. For an arbitrary number r we choose $a_{1k} = ra_{0k}$, $k = 1, \cdots, s$, and $a_{jk} = 0$, $j = 2, \cdots, n$, $k = 1, \cdots, s$. Our structure (2.1) has become

$$X_0 = (a_{01}\xi_1 + \cdots + a_{0s}\xi_s) + \eta_0$$
$$X_1 = r(a_{01}\xi_1 + \cdots + a_{0s}\xi_s) + \eta_1$$
$$X_j = \eta_j, \qquad j = 2, \cdots, n.$$

Since X_0 and X_1 are independent of $X_2, \cdots, X_n$, $E(X_0|X_1) = E(X_0|X_1, \cdots, X_n)$. We may apply corollary 2 or more simply Miss Fix's theorem to show that a necessary condition of the linearity of regression irrespective of the value of r, $-\infty < r < \infty$, is that

$$\psi_{a_{01}\xi_1+\cdots+a_{0s}\xi_s}(t) = -g(a_{01}, \cdots, a_{0s})|t|^\nu$$

$$\psi_{\eta_1}(t) = -K_1|t|^\nu$$

where $g > 0$, $K_1 > 0$, and $1 < \nu \leqq 2$.

For those values of the a_{0k} for which $a_{01}\xi_1 + \cdots + a_{0s}\xi_s$ is degenerate, $g(a_{01}, \cdots, a_{0s})$ is defined to be zero. Hence $\psi_{a_{01}\xi_1+\cdots+a_{0s}\xi_s}(t)$ being defined for all values of $a_{01}, \cdots, a_{0s}$, we have

$$\psi_{\xi_1,\cdots,\xi_s}(a_{01}, \cdots, a_{0s}) = \psi_{a_{01}\xi_1+\cdots+a_{0s}\xi_s}(1) = -g(a_{01}, \cdots, a_{0s}).$$

That g is positive homogeneous of degree ν, follows from the equalities

$$-g(ta_{01}, \cdots, ta_{0s}) = \psi_{\xi_1,\cdots,\xi_s}(ta_{01}, \cdots, ta_{0s}) = \psi_{a_{01}\xi_1+\cdots+a_{0s}\xi_s}(t)$$
$$= -|t|^\nu g(a_{01}, \cdots, a_{0s}).$$

Furthermore, it is easy to see that if any other η, say η_j, is nondegenerate, the log of its characteristic function may be written in the form given in the theorem with $K_j > 0$; but we may write the degenerate law in this form with $K_j = 0$.

It is immediately seen by an application of lemma 2 that the distribution of $\xi_1, \cdots, \xi_s$ must be symmetric multivariate stable. In particular, we have proved that if a second moment of any ξ_j, or of any one of the nondegenerate η_j is assumed to exist, then the necessary and sufficient condition for the regression of X_0 on $X_1, \cdots, X_n$ to be linear is that each η_j be normal and $\xi_1, \cdots, \xi_s$ be multivariate normal.

Now we shall find necessary and sufficient conditions for the regression to be linear. We will see that in the case $n = 1$ the conditions of theorem 2 are sufficient also, and in the case $n > 1$, except for a very singular stable distribution the normal is the only distribution that will allow a linear regression for all values of the constants.

THEOREM 3. *In the case $n = 1$, a necessary and sufficient condition that the regression of X_0 on X_1 be linear irrespective of the values of the constants a_{jk}, is that*

$$\psi_{\xi_1,\cdots,\xi_s}(t_1, \cdots, t_s) = -g(t_1, \cdots, t_s)$$

$$\psi_{\eta_1}(t) = -K|t|^\nu,$$

where $g(t_1, \cdots, t_s)$ is a positive homogeneous real function of degree ν, $1 < \nu \leqq 2$, and $K > 0$.

Proof. The necessity was proved in thereom 2. To prove sufficiency it must be shown that for all values of the a_{jk}, there exists a solution for b independent of the value of t in equation (3) of lemma 1 with $n = 1$.

$$\sum_{k=1}^{s} \psi_{\xi_1,\cdots,\xi_k',\cdots,\xi_s}(ta_{11}, \cdots, ta_{1s})(a_{0k} - ba_{1k}) = b\psi_{\eta_1}'(t). \tag{4.1}$$

Since $\xi_1, \cdots, \xi_s$ all possess finite first moments, all first partial derivatives of $g(t_1, \cdots, t_s)$ must exist.

$$\psi_{\xi_1,\cdots,\xi_k',\cdots,\xi_s}(t_1, \cdots, t_s) = -g_k(t_1, \cdots, t_s)$$

$$\psi_{\eta_1}'(t) = -K\nu\frac{|t|^\nu}{t},$$

where $g_k(t_1, \cdots, t_s) = (\partial/\partial t_k)g(t_1, \cdots, t_s)$. Substitution into equation (4.1) yields

$$\sum_{k=1}^{s} [-g_k(ta_{11}, \cdots, ta_{1s})](a_{0k} - ba_{1k}) = b\left[-K\nu\frac{|t|^\nu}{t}\right]. \tag{4.2}$$

For differentiable positively homogeneous functions of degree ν we have

$$g_k(tu_1, \cdots, tu_s) = \frac{|t|^\nu}{t} g_k(u_1, \cdots, u_s). \tag{4.3}$$

This is true since

$$|t|^\nu g_k(u_1, \cdots, u_s) = \frac{\partial}{\partial u_k}|t|^\nu g(u_1, \cdots, u_s) = \frac{\partial}{\partial u_k} g(tu_1, \cdots, tu_s)$$
$$= tg_k(tu_1, \cdots, tu_s).$$

Thus $-|t|^\nu/t$ may be canceled from both sides of equation (4.2).

$$\sum_{k=1}^{s} g_k(a_{11}, \cdots, a_{1s})a_{0k} = b\left[\sum_{k=1}^{s} a_{1k}\frac{\partial}{\partial a_{1k}} g(a_{11}, \cdots, a_{1s}) + K\nu\right].$$

It remains to be shown that the coefficient of b is never zero. This is done using Euler's theorem for a homogeneous function of degree ν, that is,

$$\sum_{k=1}^{s} a_{1k} \frac{\partial}{\partial a_{1k}} g(a_{11}, \cdots, a_{1s}) = \nu g(a_{11}, \cdots, a_{1s}),$$

so that the coefficient of b is $\nu [g(a_{11}, \cdots, a_{1s}) + K]$. This is positive since $g \geqq 0$ by a well-known property of characteristic functions and $K > 0$ and $\nu > 1$ by hypothesis.

THEOREM 4. *In the case $n > 1$, a necessary and sufficient condition for the regression of X_0 on $X_1, \cdots, X_n$ to be linear in $X_1, \cdots, X_n$, irrespective of the values of the constants a_{jk} is that either*

(i) *both $\xi_1, \cdots, \xi_s$ be multivariate normal and η_1 through η_n be normal, or*

(ii) *η_2 through η_n be degenerate and*

$$\psi_{\xi_1, \cdots, \xi_s}(t_1, \cdots, t_s) = -|k_1 t_1 + \cdots + k_s t_s|^\nu$$

$$\psi_{\eta_1}(t) = -K|t|^\nu,$$

where $1 < \nu < 2$, $k_1 \neq 0, \cdots, k_s \neq 0$, and $K > 0$.

Remark. The condition that $\psi_{\xi_1, \cdots \xi_s}(t_1, \cdots, t_s) = -|k_1 t_1 + \cdots + k_s t_s|^\nu$, $k_j \neq 0$, $j = 1, \cdots, s$ may be replaced by the condition that each ξ_k be a nonzero multiple of a fixed symmetric stable random variable ξ with characteristic function $\phi_\xi(t) = \exp(-|t|^\nu)$. This follows since

$$\phi_{\xi_1, \cdots, \xi_s}(t_1, \cdots, t_s) = \left[\exp - \left|\sum_{j=1}^{s} k_j t_j\right|^\nu\right] = \phi_\xi(\sum k_j t_j) = E(\exp i \sum t_j k_j \xi)$$

$$= \phi_{k_1\xi, \cdots, k_s\xi}(t_1, \cdots, t_s)$$

and conversely if each $\xi_j = k_j \xi$ then

$$\phi_{\xi_1, \cdots, \xi_s}(t_1, \cdots, t_s) = E\left[\exp i \sum t_j k_j \xi\right] = \exp - \left|\sum t_j k_j\right|^\nu.$$

Proof of necessity. We begin where theorem 2 leaves off. We must show that either $\nu = 2$ in that theorem in which case $\xi_1, \cdots, \xi_s$ will be multivariate normal from lemma 2, or if $1 < \nu < 2$, then $K_j = 0$ for $j = 2, \cdots, n$ and $g(t_1, \cdots, t_s) = |k_1 t_1 + \cdots + k_s t_s|^\nu$.

First it will be proved that if some K_j for $j \geqq 2$ is positive, then ν must be equal to two. Suppose $K_2 > 0$. With $a_{01} = a_{11} = a_{21} = 1$ and all the rest of the $a_{jk} = 0$, equation (3) of lemma 1 becomes the following necessary condition for linear regression

$$\psi'_{\xi_1}(t_1 + t_2)(1 - b_1 - b_2) = \sum_{h=1}^{n} b_h \psi'_{\eta}(t_h).$$

There must exist a solution for the b_j, $j = 2, \cdots, n$ which is independent of the t_j, $j = 1, 2, \cdots, n$. So fix $t_j = 0$ for $j = 3, 4, \cdots, n$

$$\psi'_{\xi_1}(t_1 + t_2)(1 - b_1 - b_2) = b_1 \psi'_{\eta_1}(t_1) + b_2 \psi'_{\eta_2}(t_2). \tag{4.4}$$

Putting alternately $t_1 = 0$ and $t_2 = 0$, we must have

$$\psi'_{\xi_1}(t_1 + t_2)(1 - b_1 - b_2) = \psi'_{\xi_1}(t_1)(1 - b_1 - b_2) + \psi'_{\xi_1}(t_2)(1 - b_1 - b_2).$$

Suppose $1 - b_1 - b_2 = 0$; then equation (4.4) because of the nondegeneracy of both η_1 and η_2 would imply that $b_1 = 0$ and $b_2 = 0$. Hence $1 - b_1 - b_2 = 1$. This contradiction permits us to cancel $(1 - b_1 - b_2)$.

$$\psi'_{\xi_1}(t_1 + t_2) = \psi'_{\xi_1}(t_1) + \psi'_{\xi_1}(t_2).$$

Since $\psi'_{\xi_1}(t)$ must be continuous, this equation implies that $\psi'_{\xi_1}(t)$ is a linear function of t and hence that $\psi_{\xi_1}(t)$ is a quadratic. Thus ξ_1 and hence η_1 and η_2 must be normal, i.e., $\nu = 2$.

Hence we know that if $n > 1$ and $1 < \nu < 2$, then $K_j = 0$ for $j = 2, \cdots, n$. We must show that in this case it is also necessary that $g(t_1, \cdots, t_s) = |k_1 t_1 + \cdots + k_s t_s|^\nu$. The result is obvious for $s = 1$ since g is positive homogeneous of degree ν. It will next be proved for $s = 2$, and then extended to arbitrary s.

Assume $s = 2$. Since $K_j = 0$, $j = 2, \cdots, n$, it is necessary by theorem 2 that

$$\psi_{\xi_1, \xi_2}(t_1, t_2) = -g(t_1, t_2)$$

$$\psi_{\eta_1}(t) = -K_1|t|^\nu$$

$$\psi_{\eta_j}(t) = 0, \qquad j = 2, \cdots, n.$$

We may take derivatives with respect to t_1 and t_2, and put them in condition (3) of lemma 1, with $a_{jk} = 0$ for $j = 3, \cdots, n$ and all $k = 1, 2$.

$$g_1(t_1 a_{11} + t_2 a_{21}, t_1 a_{12} + t_2 a_{22})(a_{01} - b_1 a_{11} - b_2 a_{21}) + g_2(t_1 a_{11} + t_2 a_{21}, t_1 a_{12} + t_2 a_{22})(a_{02} - b_1 a_{12} - b_2 a_{22}) = b_1 K_1 \nu \frac{|t_1|^\nu}{t_1}. \tag{4.5}$$

Restrict the values of the a_{jk}'s further so that $a_{21} \neq 0$, $a_{22} \neq 0$ and

$$\Delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \neq 0.$$

Let

$$V_1 = t_1 a_{11} + t_2 a_{21}$$

$$V_2 = t_1 a_{12} + t_2 a_{22}$$

then

$$t_1 = \frac{1}{\Delta}[a_{22}V_1 - a_{21}V_2]$$

$$t_2 = \frac{1}{\Delta}[a_{11}V_2 - a_{12}V_1].$$

Equation (4.5) becomes

$$g_1(V_1, V_2)\alpha_1 + g_2(V_1, V_2)\alpha_2 = \frac{|a_{22}V_1 - a_{21}V_2|^\nu}{a_{22}V_1 - a_{21}V_2} b_1 , \tag{4.6}$$

where

$$\alpha_1 = \frac{1}{K_{1\nu}} \frac{|\Delta|^\nu}{\Delta} (a_{01} - b_1 a_{11} - b_2 a_{21})$$
$$\alpha_2 = \frac{1}{K_{1\nu}} \frac{|\Delta|^\nu}{\Delta} (a_{02} - b_1 a_{12} - b_2 a_{22}) . \tag{4.7}$$

Since $\Delta \neq 0$, taking arbitrary t_1 and t_2 is equivalent to taking arbitrary V_1 and V_2.

$$\text{Take } V_1 = 0 \text{ and } V_2 = 1; \text{then } g_1(0, 1)\alpha_1 + g_2(0, 1)\alpha_2 = -\frac{|a_{21}|^\nu}{a_{21}} b_1.$$
$$\text{Take } V_1 = 1 \text{ and } V_2 = 0; \text{then } g_1(1, 0)\alpha_1 + g_2(1, 0)\alpha_2 = \frac{|a_{22}|^\nu}{a_{22}} b_1. \tag{4.8}$$

Let

$$\Delta_1 = \begin{vmatrix} g_1(0, 1) & g_2(0, 1) \\ g_1(1, 0) & g_2(1, 0) \end{vmatrix} = g_1(0, 1)g_2(1, 0) - g_2(0, 1)g_1(1, 0) .$$

It will now be shown that $\Delta_1 = 0$. Suppose $\Delta_1 \neq 0$, then solving equations (4.8)

$$\alpha_1 = -\frac{b_1}{\Delta_1}\left[g_2(1, 0) \frac{|a_{21}|^\nu}{a_{21}} + g_2(0, 1) \frac{|a_{22}|^\nu}{a_{22}} \right]$$
$$\alpha_2 = \frac{b_1}{\Delta_1}\left[g_1(1, 0) \frac{|a_{21}|^\nu}{a_{21}} + g_1(0, 1) \frac{|a_{22}|^\nu}{a_{22}} \right] . \tag{4.9}$$

If $b_1 = 0$ then $\alpha_1 = \alpha_2 = 0$, so that from (4.7), $a_{01} = b_2 a_{21}$ and $a_{02} = b_2 a_{22}$. Hence by making the additional restriction that

$$\begin{vmatrix} a_{01} & a_{02} \\ a_{21} & a_{22} \end{vmatrix} \neq 0 ,$$

it is necessary that $b_1 \neq 0$.

Equations (4.9) may be substituted into equation (4.6) with $V_1 = V_2 = 1$, and b_1 may be canceled.

$$[g_2(1, 1)g_1(0, 1) - g_1(1, 1)g_2(0, 1)] \frac{|a_{22}|^\nu}{a_{22}} - [g_1(1, 1)g_2(1, 0) - g_2(1, 1)g_1(1, 0)] \frac{|a_{21}|^\nu}{a_{21}} \equiv \Delta_1 \frac{|a_{22} - a_{21}|^\nu}{a_{22} - a_{21}} , \tag{4.10}$$

the identity holding in all values of a_{22} and a_{21} different from zero. Take successively the limit as $a_{21} \to 0$ and as $a_{22} \to 0$.

$$g_2(1, 1)g_1(0, 1) - g_1(1, 1)g_2(0, 1) = \Delta_1$$
$$g_1(1, 1)g_2(1, 0) - g_2(1, 1)g_1(1, 0) = \Delta_1 .$$

Equation (4.10) with $\Delta_1 \neq 0$ canceled becomes

$$\frac{|a_{22}|^\nu}{a_{22}} - \frac{|a_{21}|^\nu}{a_{21}} \equiv \frac{|a_{22} - a_{21}|^\nu}{a_{22} - a_{21}} .$$

This can happen only if $\nu = 2$; contradiction. Hence $\Delta_1 = 0$.

It will now be shown that it is necessary that $b_1 = 0$. Since $g_1(1, 0) = d/dt \log \phi_{\xi_1}(t)|_{t=1}$ and ξ_1 is stable, the nondegeneracy of ξ_1 implies that $g_1(1, 0) \neq 0$; similarly, $g_2(0, 1) \neq 0$. Consequently, since $\Delta_1 = 0$, $g_1(0, 1) \neq 0$ and $g_2(1, 0) \neq 0$. Therefore, let

$$r = \frac{g_1(0, 1)}{g_2(0, 1)} = \frac{g_1(1, 0)}{g_2(1, 0)} .$$

Equations (4.8) become

$$g_2(0, 1)[r\alpha_1 + \alpha_2] = -\frac{|a_{21}|^\nu}{a_{21}} b_1$$
$$g_2(1, 0)[r\alpha_1 + \alpha_2] = \frac{|a_{22}|^\nu}{a_{22}} b_1 .$$

Therefore,

$$-g_2(1, 0) \frac{|a_{21}|^\nu}{a_{21}} b_1 = g_2(0, 1) \frac{|a_{22}|^\nu}{a_{22}} b_1 .$$

Hence, except possibly for those values of a_{21} and a_{22} such that

$$\frac{|a_{21}|^\nu}{a_{21}} g_2(1, 0) + \frac{|a_{22}|^\nu}{a_{22}} g_2(0, 1) = 0 ,$$

b_1 must be equal to zero.

Thus the necessary condition (4.6) becomes

$$g_1(V_1, V_2)\alpha_1 + g_2(V_1, V_2)\alpha_2 \underset{V_1V_2}{\equiv} 0 , \tag{4.11}$$

where

$$\alpha_1 = \frac{1}{K_{1\nu}} \frac{|\Delta|^\nu}{\Delta} (a_{01} - b_2 a_{21})$$
$$\alpha_2 = \frac{1}{K_{1\nu}} \frac{|\Delta|^\nu}{\Delta} (a_{02} - b_2 a_{22}) .$$

In view of the restriction

$$\begin{vmatrix} a_{01} & a_{02} \\ a_{21} & a_{22} \end{vmatrix} \neq 0\,,$$

not both α_1 and α_2 can be zero. It is necessary that g be differentiable and so we are looking for the most general differentiable solution of the partial differential equation (4.11). An obvious solution is $g(V_1, V_2) = \alpha_2 V_1 - \alpha_1 V_2$. Hence a large class of solutions is given by $g(V_1, V_2) = f(\alpha_2 V_1 - \alpha_1 V_2)$ where f is an arbitrary differentiable function. To see that this is the most general solution, let $f(V_1, V_2)$ be any solution. Suppose without loss of generality that $\alpha_1 \neq 0$. Then let

$$u_1 = \alpha_2 V_1 - \alpha_1 V_2$$
$$u_2 = V_1.$$

The determinant is not zero, so we may write f as a function of u_1 and u_2, $f(V_1, V_2) = \bar{f}(u_1, u_2)$

$$\frac{\partial}{\partial V_1}\bar{f}(u_1, u_2) = \alpha_2 \bar{f}_1(u_1, u_2) + \bar{f}_2(u_1, u_2)$$

$$\frac{\partial}{\partial V_2}\bar{f}(u_1, u_2) = -\alpha_1 \bar{f}_1(u_1, u_2)\,.$$

Substituting these derivatives into equation (4.11) we see that $\bar{f}_2(u_1, u_2) = 0$; that is, that f is a function of $u_1 = \alpha_2 V_1 - \alpha_1 V_2$ alone. We also know that g is positive homogeneous of degree ν. From these two facts it is easily seen that $g(V_1, V_2) = |k_1 V_1 + k_2 V_2|^{\nu}$ where $k_1 \neq 0$ and $k_2 \neq 0$ from our assumption of nondegeneracy.

From the remark made after the statement of the theorem, this is equivalent to the condition $\xi_2 = k\xi_1$, where ξ_1 is nondegenerate stable and $k \neq 0$ is a constant.

Now let s be arbitrary, $s > 2$. Putting all the $a_{jk} = 0$ for $k = 3, \cdots, s, j = 0, 1, \cdots, n$, we are reduced to the case $s = 2$; hence $\xi_2 = c_2\xi_1$ with ξ_1 stable. By symmetry we must also have $\xi_k = c_k\xi_1$ for $k = 3, \cdots, s$, $c_k \neq 0$. Again we may apply the remark just mentioned to show that

$$\psi_{\xi_1,\cdots,\xi_s}(t_1, \cdots, t_s) = -\left|\sum_{j=1}^{s} t_j k_j\right|^{\nu},$$

which completes the proof of necessity.

Proof of sufficiency. As mentioned before, that the regression is linear when all variables are normal is well known. We concern ourselves with the case $1 < \nu < 2$. In this case

$$\psi_{\xi_1,\cdots,\xi_j',\cdots,\xi_s}(t_1, \cdots, t_s) = -k_j\nu \frac{\left|\sum_{i=1}^{s} k_i t_i\right|^{\nu}}{\sum_{i=1}^{s} k_i t_i},$$

$$k_i \neq 0\,, \qquad i = 1, \cdots, s$$

$$\psi_{\eta_1}'(t) = -K_1\nu \frac{|t|^{\nu}}{t}\,, \qquad K_1 > 0$$

$$\psi_{\eta_j}'(t) = 0\,, \qquad \text{for } j = 2, \cdots, n$$

so that condition (3) of lemma 1 becomes

$$(4.12)\qquad \sum_{l=1}^{s}\left[k_l \frac{\left|\sum_{i=1}^{s} k_i \sum_{j=1}^{n} t_j a_{ji}\right|^{\nu}}{\sum_{i=1}^{s} k_i \sum_{j=1}^{n} t_j a_{ji}}\right]\left(a_{0l} - \sum_{h=1}^{n} b_h a_{hl}\right) = b_1 K_1 \frac{|t_1|^{\nu}}{t_1}\,.$$

It is sufficient to show that for all values of the a_{jk} there exist functions $b_1, \cdots, b_n$ of them such that (4.12) is an identity in $t_1, \cdots, t_n$.

Unless $\sum_{l=1}^{s} a_{hl}k_l = 0$ for $h = 2, \cdots, n$, we may choose $b_1 = 0$ and $b_2, \cdots, b_n$ so that $\sum_{h=2}^{n} b_h \sum_{l=1}^{s} a_{hl}k_l = \sum_{l=1}^{s} k_l a_{0l}$. This will satisfy (4.12) identically in $t_1, \cdots, t_n$. If, however, $\sum_{l=1}^{s} a_{hl}k_l = 0$ for $h = 2, \cdots, n$, then equation (4.12) becomes

$$\frac{|t_1|^{\nu}}{t_1}\frac{\left|\sum_{l=1}^{s} k_l a_{1l}\right|^{\nu}}{\sum_{l=1}^{s} k_l a_{1l}}\left[\sum_{l=1}^{s} k_l a_{0l} - b_1 \sum_{l=1}^{s} k_l a_{1l}\right] = K_1 b_1 \frac{|t_1|^{\nu}}{t_1}\,.$$

Canceling $|t_1|^{\nu}/t_1$ and solving for b_1, we have

$$b_1\left[K_1 + \left|\sum_{l=1}^{s} k_l a_{1l}\right|^{\nu}\right] = \frac{\left|\sum_{l=1}^{s} k_l a_{1l}\right|^{\nu}}{\sum_{l=1}^{s} k_l a_{1l}} \sum_{l=1}^{s} k_l a_{0l}\,.$$

A solution of the equation for b_1 always exists since

$$K_1 + \left|\sum_{l=1}^{s} k_l a_{1l}\right|^{\nu} > 0\,.$$

5. A characterization of the normal law. If the linear structural relation of section 2 is specialized to the case $s = 1$, $n \geqq 2$, a much stronger result than any of those found in the previous sections may be proved. The unique position reserved for the normal distribution in probability theory is due historically to the central limit theorem. But in recent times this distribution has appeared again and again in connection with various and diverse problems. Particularly interesting in this respect are the theorems of R. C. Geary [**5**], Bernstein-Basu-Darmois [**3**], and E. Lukacs [**14**]. Theorems in this paper and particularly the following one may be added to this list of characterizations of the normal law.

Note that in theorem 5, the regression is assumed to be linear for only one set of values of the constants a_j, and no assumption of existence of moments other than the first is made. This theorem has been independently proved by R. G. Laha [**10**] of the Indian Statistical Institute.

THEOREM 5. *Let ξ, η_1, η_2, $\cdots$, η_n, $n \geqq 2$, be independent, nondegenerate random* *variables with finite first moments. Let* $X_j = a_j\xi + \eta_j$ *with* $a_j \neq 0, j = 1, 2, \cdots, n$. *In order that the multiple regression of* ξ *on* $X_1, \cdots, X_n$ *be linear, it is necessary and sufficient that* $\xi, \eta_1, \cdots, \eta_n$ *be normal.*

Proof. Assume without loss of generality that all the first moments are zero. Condition (3) of lemma 1 gives

$$\psi_\xi'\left(\sum_{j=1}^n t_j a_j\right)\left(1 - \sum_{h=1}^n b_h a_h\right) = \sum_{j=1}^n b_j \psi_{\eta_j}'(t_j) . \tag{5.1}$$

We will show that ξ and η_1 are normal, and by symmetry, η_2 through η_n will be normal. In (5.1) put $t_3 = t_4 = \cdots t_n = 0$.

$$\psi_\xi'(t_1a_1 + t_2a_2)\left(1 - \sum_1^n b_h a_h\right) = b_1\psi_{\eta_1}'(t_1) + b_2\psi_{\eta_2}'(t_2) .$$

Putting successively $t_1 = 0$ and $t_2 = 0$, and substituting the result back into the equation, we have

$$\psi_\xi'(t_1a_1 + t_2a_2)\left(1 - \sum_1^n b_h a_h\right) = [\psi_\xi'(t_1a_1) + \psi_\xi'(t_2a_2)]\left(1 - \sum_1^n b_h a_h\right) .$$

If $1 - \sum b_h a_h = 0$, then by (5.1) the nondegeneracy of $\eta_1, \cdots, \eta_n$ would imply $b_j = 0, j = 1, \cdots, n$, so that $1 - \sum b_h a_h = 1$. Hence we may cancel $1 - \sum b_h a_h$. The function $\psi_\xi'(t)$ being continuous is then seen to be a linear function of t, and hence $\psi_\xi(t)$ is a quadratic at least in a neighborhood of the origin. From this it is well known that ξ and hence η_1 and η_2 must be normal.

The difficulties encountered in trying to extend this result to the general case where s is arbitrary may be well illustrated in the case $n = 2, s = 2$. In fact, it is sufficient to consider the case where ξ_1 and ξ_2 are independent, prefaced with the remark that even worse anomalies may occur if ξ_1 and ξ_2 are allowed to be dependent. In the case considered, condition (3) of lemma 1 becomes

$$\gamma_1\psi_{\xi_1}'(t_1a_{11} + t_2a_{21}) + \gamma_2\psi_{\xi_2}'(t_1a_{12} + t_2a_{22}) = b_1\psi_{\eta_1}'(t_1) + b_2\psi_{\eta_2}'(t_2) , \tag{5.2}$$

where $\gamma_1 = (a_{01} - b_1a_{11} - b_2a_{21})$ and $\gamma_2 = (a_{02} - b_1a_{12} - b_2a_{22})$.

First, it might happen that $b_1 = 0$, in which case no restriction would be placed on the variable η_1. It may be shown, however, that whenever $b_1 \neq 0$, η_1 must be a normal variable. Similarly for η_2.

Second, it may happen that $\gamma_1 = 0$, in which case no restriction would be placed on the variable ξ_1, but in this case it may be shown that ξ_2, η_1, and η_2 must be normal variables.

Finally, if $\Delta = a_{11}a_{22} - a_{12}a_{21} \neq 0$ and $b_1 \neq 0$, $b_2 \neq 0$, $\gamma_1 \neq 0$ and $\gamma_2 \neq 0$, all variables concerned must be normal. However, if $\Delta = 0$, then ψ_{ξ_1}' and ψ_{ξ_2}' have essentially the same argument, and thus the distribution of ξ_1 may be "subtracted out" if ξ_2 has the same distribution as a component. It is easily checked that the following example satisfies equation (5.2).

Let Φ *represent a normal distribution with zero mean and unit variance, and let* F

represent any distribution with a zero mean. Let $\xi_1, \xi_2, \eta_1, \eta_2$ *have respectively the distributions given by* $F, F * \Phi, \Phi, \Phi$. *Then for the structure*

$$X_0 = \xi_1 + 3\xi_2$$

$$X_1 = \xi_1 + \xi_2 + \eta_1$$

$$X_2 = \xi_1 + \xi_2 + \eta_2$$

we have $E(X_0 | X_1, X_2) = X_1 + X_2$.

If $s > 2$, an example similar to this one may be given, in which none of the determinants, such as Δ above, are zero.

6. Dependent error terms. First we consider a theorem on an extension of the problem of Ragnar Frisch allowing the error terms to have a bivariate stable distribution. Secondly, we consider the implications of this and other theorems in this paper on the problem of identifiability in a linear structural relation.

In the simple structural relation (1.1), we assume that ξ is independent of (η_0, η_1) and that (η_0, η_1) has a bivariate stable distribution which, for simplicity, we take to be symmetric. We assume further that all of the variables have finite first moments which we take to be zero. Thus, the characteristic function of (η_0, η_1) may be written in the form

$$\phi_{\eta_0,\eta_1}(t_0, t_1) = e^{-g(t_0,t_1)} ,$$

where g is a real positive homogeneous function of degree ν, $1 < \nu \leqq 2$. To avoid cases where the regression is trivially linear we assume also that η_1 is nondegenerate and $a_1 \neq 0$.

THEOREM 6. *In order that the regression of* X_0 *on* X_1 *be linear, it is necessary and sufficient that either*

(i) ξ *is symmetric stable with the same characteristic exponent* ν, *(perhaps degenerate), or*

$$\text{(ii)} \quad \frac{a_0}{a_1} = \frac{g_0(0, 1)}{g_1(0, 1)} .$$

Proof of necessity. Using condition (2) of lemma 1 and the particular form of the characteristic function of (η_0, η_1), we may derive a necessary and sufficient condition similar to (3).

$$\psi_\xi'(ta_1)(a_0 - ba_1) = g_0(0, t) - bg_1(0, t) ,$$

which holds for all values of t. Using formula (4.3), this equation may be rewritten as

$$\psi_\xi'(ta_1)(a_0 - ba_1) = -[bg_1(0, 1) - g_0(0, 1)]\frac{|t|^\nu}{t} . \tag{6.1}$$

Hence, if $a_0 - ba_1 \neq 0$, ξ must be symmetric stable with characteristic exponent ν. If $a_0 - ba_1 = 0$, then also $bg_1(0, 1) - g_0(0, 1) = 0$.

Proof of sufficiency. The sufficiency of condition (ii) is obvious from equation (6.1). Therefore, let $\psi_\xi(t) = -K|t|^\nu$, $K \geqq 0$. Equation (6.1) becomes

$$-\nu K \frac{|t|^\nu}{t}\frac{|a_1|^\nu}{a_1}(a_0 - ba_1) = -[bg_1(0, 1) - g_0(0, 1)]\frac{|t|^\nu}{t} .$$

Upon rearranging we find

$$b[\nu K|a_1|^\nu + g_1(0, 1)] = \nu K \frac{|a_1|^\nu}{a_1} a_0 - g_0(0, 1) .$$

The coefficient of b is not zero since η_1 is nondegenerate stable. Hence, a solution for b always exists.

In particular this theorem implies that if the regression is linear for two values of the pair of constants (a_0, a_1) and (a_0', a_1') such that $a_0/a_1 \neq a_0'/a_1'$, then the variable ξ must be symmetric stable.

7. Relation to identifiability. Particularly interesting is the relation between these theorems and the theorems on identifiability in a linear structural relation in a paper by Reiersøl [**17**], which also contains an interesting account of the history of the subject. In a structure such as (1) we are able to observe only the variables X_0 and X_1, and by taking an infinite number of observations we would be able almost surely to construct their joint distribution by the Glevenko-Cantelli theorem. However, we are usually interested in the latent apparatus of the structure. We would like to know the hidden relation between X_0 and X_1: the ratio a_1/a_0 and the distributions of ξ, η_0 and η_1. We shall rewrite equation (1.1) in the form used by Reiersøl.

$$X_0 = \xi + \eta_0$$

$$X_1 = \beta\xi + \beta_0 + \eta_1 ,$$

where β replaces the ratio a_1/a_0 and β_0 is some arbitrary constant.

In his paper Reiersøl considers two models for this structure. Essentially, model A is composed of the three assumptions:

(1) ξ is independent of (η_0, η_1) ,

(2) $E\eta_0 = E\eta_1 = 0$,

(3) η_0 and η_1 are jointly normal.

Model B is composed of assumptions (1) and (2) of model A but with (3) replaced by (3′) η_0 and η_1 are independent.

The parameter β (or β_0 or the distributions of ξ, η_0 and η_1) is said to be identifiable in a model if the joint distribution of X_0, X_1 determines β (or β_0 or the distributions of ξ, η_0 and η_1) uniquely—that is, if there do not exist two different realizations of the model, $(\beta, \beta_0, \xi, \eta_0, \eta_1)$ and $(\beta^*, \beta_0^*, \xi^*, \eta_0^*, \eta_1^*)$ with $\beta \neq \beta^*$, which give rise to the same joint distribution of X_0, X_1. Thus, if there exists an estimate of the parameter β, which is consistent in the model, β is identifiable in the model. That the converse is true in many important cases, in particular in models A and B above, has been demonstrated by J. Neyman [**15**] by actually constructing consistent estimates of β. Interesting papers on identifiability in many dimensional structural relations have been written by Koopmans, Rubin, and Leipnik [**9**], and T. A. Jeeves [**7**].

Reiersøl establishes necessary and sufficient conditions for β to be identifiable in each of the models A and B.

In model A, β *is not identifiable if and only if* ξ *is normal.*

In model B, β *is not identifiable if and only if* ξ *is normal and either the distribution of* η_0 *or the distribution of* η_1 *has a normal distribution as a component.*

These theorems stand in an interesting juxtaposition to the previous theorems of this paper. They tend to indicate that one cannot simultaneously assume that the regression of X_0 on X_1 is linear and expect to be able to find a consistent estimate of β. For example, in model A, if the regression of X_0 on X_1 is linear, then by theorem 6, β is identifiable only if $\beta = \sigma_1/\rho$ where σ_1^2 is the variance of η_1 and ρ is the correlation coefficient of (η_0, η_1). Also in model B, if we assume the regression is linear whatever be the value of β in some interval or even linear for just two values of β, and we assume that a second moment of ξ or η_1 exists, then an application of corollary 1 of section 3 will tell us that β is not identifiable.

The correspondence between theorem 6 and Reiersøl's theorem for model A can be strengthened by a slight extension of Reiersøl's result. We may enlarge model A by replacing condition (3) by

(3″) η_0 and η_1 are nonsingular, symmetric bivariate stable.

Denote the model composed of assumptions (1), (2) and (3″) as model A_1. This model has an independent interest in that it is an example of a model for which ξ may not be normal and yet β will still not be identifiable.

THEOREM 7. *In model* A_1, β *is not identifiable if and only if* ξ *is symmetric stable with the same characteristic exponent as* (η_0, η_1).

By symmetric stable we do not mean here that ξ is symmetric about zero but rather symmetric about some point. We shall omit the proof since it follows the lines of the proof of Reiersøl's theorem rather closely. It is also easy to show that the theorem is still true if the word symmetric is omitted both from the model and the theorem.

It is of some further interest to consider identifiability in a model analogous to that of theorem 5. Although the result of theorem 8 below does not parallel that of theorem 5 quite so nicely, its interest lies in the fact that it complements Reiersøl's result and affords an example where all the variables concerned are normal and still we may have identifiability.

Suppose random variables X_j have a structure of the form

$$X_0 = \xi + \eta_0$$

$$X_j = a_j\xi + \eta_j , \qquad j = 1, 2, \cdots, n , \qquad n \geqq 2$$

We shall use a model composed of the following three assumptions.

(1) ξ is independent of $\eta_0, \eta_1, \cdots, \eta_n$,

(2) $E\eta_0 = E\eta_1 = \cdots = E\eta_n = 0$,

(3) $\eta_0, \eta_1, \cdots, \eta_n$ are completely independent.

We are interested in the identifiability of one of the a_j, say a_1. But if $a_2 = a_3 = \cdots = a_n = 0$, then the problem is reduced to Reiersøl's theorem for model B. If we assume $a_2 \neq 0$ we obtain the following theorem.

THEOREM 8. *If* $a_2 \neq 0$, a_1 *is not identifiable in the model if and only if* ξ *is degenerate.*

Proof. The sufficiency of degeneracy is obvious. To prove necessity first note that from Reiersøl's theorem it is necessary that ξ be normal. We are given the joint distribution of X_0, X_1, X_2; hence, we know the value of

$$\psi_{X_0,X_1,X_2}(u_0, u_1, u_2) = \psi_\xi(u_0 + a_1u_1 + a_2u_2) + \psi_{\eta_0}(u_0) + \psi_{\eta_1}(u_1) + \psi_{\eta_2}(u_2)$$

in a neighborhood of the origin. If there is a different realization of the model $(a_1^*, a_2^*, \xi^*, \eta_0^*, \eta_1^*, \eta_2^*)$ with $a_1 \neq a_1^*$, yielding the same distribution of X_0, X_1, X_2 we must have

$$im(u_0 + a_1u_1 + a_2u_2) - \frac{K^2}{2}(u_0 + a_1u_1 + a_2u_2)^2 + \psi_{\eta_0}(u_0) + \psi_{\eta_1}(u_1) + \psi_{\eta_2}(u_2)$$

$$= im^*(u_0+a_1^*u_1+a_2^*u_2) - \frac{K^{*2}}{2}(u_0+a_1^*u_1+a_2^*u_2)+\psi_{\eta_0^*}(u_0)+\psi_{\eta_1^*}(u_1)+\psi_{\eta_2^*}(u_2),$$

where m and K^2 (similarly m^* and K^{*2}) are the mean and the variance of ξ (respectively ξ^*). Applying the partial differential operators $\partial^2/\partial u_0\partial u_2$ and $\partial^2/\partial u_1\partial u_2$ in turn, we derive the equalities

$$Ka_2 = K^* a_2^*$$

$$Ka_1a_2 = K^* a_1^* a_2^* .$$

From this follows the equation

$$Ka_1a_2 = Ka_1^* a_2 .$$

Since $a_1 \neq a_1^*$ and $a_2 \neq 0$, we must have $K = 0$; i.e., ξ is degenerate.

I should like to express my gratitude and indebtedness to Dr. Lucien LeCam for the suggestion of the problem here considered and for his generous advice and encouragement.

REFERENCES

[1] H. V. Allen, "A theorem concerning the linearity of regression," *Stat. Res. Mem.*, Vol. 2 (1938), pp. 60–68.

[2] Harald Cramer, *Mathematical Methods of Statistics*, Princeton, Princeton University Press, 1946.

[3] George Darmois, "Analyses générale des liasons stochastiques," *Revue de l'Institute International de Statistique*, Vol. 21 (1953), pp. 2–8.

[4] Evelyn Fix, "Distributions which lead to linear regressions," *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, University of California Press, 1949, pp. 79–91.

[5] R. C. Geary, "Distribution of Student's ratio for non-normal samples," *Roy. Stat. Soc. Jour., Supp.*, Vol. 3, (1936), pp. 178–184.

[6] B. V. Gnedenko and A. N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables*, Cambridge, Mass., Addison-Wesley Pub. Co., 1954.

[7] T. A. Jeeves, "Identification and estimation of linear manifolds in n-dimensions," *Ann. Math. Stat.*, Vol. 25 (1954), pp. 714–723.

[8] J. F. Kenney, "The regression systems of two sums having random elements in common," *Ann. Math. Stat.*, Vol. 10 (1939), pp. 70–73.

[9] T. C. Koopmans, H. Rubin and R. B. Leipnik, "Measuring the equation systems of dynamic economics," in Cowles Commission, *Statistical Inference in Dynamic Economic Models*, New Yoak, Wiley, 1950.

[10] R. G. Laha, to be published.

[11] Paul Lévy, *Théorie de l'addition des variables aleatoires*, Paris, Gauthier-Villars, 1937.

[12] Paul Lévy, *Processus stochastiques et mouvement brownien*, Paris, Gauthier-Villars, 1948.

[13] D. V. Lindley, "Regression lines and the linear structural relationship," *Roy. Stat. Soc. Jour., Supp.* Vol. 9, (1947), pp. 218–244.

[14] Eugene Lukacs, "Sur une propriété de la loi de Gauss-Laplace," *C. R. Acad. Sci. Paris*, Vol. 238 (1954), pp. 444–446.

[15] J. Neyman, "Existence of consistent estimates of the directional parameter in a linear structural relation between two variables," *Ann. Math. Stat.*, Vol. 22 (1951), pp. 497–512.

[16] C. R. Rao, "Note on a problem of Ragnar Frisch," *Econometrica* Vol. 15 (1947), pp. 245–250.

[17] Olav Reiersøl, "Identifiability of a linear relationship between variables which are subject to error," *Econometrica*, Vol. 18 (1950), pp. 375–389.