Stability and Separability

The Role of the Stable Distributions in Portfolio Theory and Some Implications for Multivariate Statistical Analysis

by

Stephen A. Ross

Working Paper No. 8-76

The author is grateful to the Rodney White Foundation, the Guggenheim Foundation and to the National Science Foundation, Grant #SOC74-20292 A01 for their aid.

July 1976

The stable distributions have a long and somewhat checkered career in finance. They were first introduced by Mandelbrot who used them to study speculative price series. On the basis of statistical criteria, the normal distribution did not appear to provide enough weight in the tails to explain the outlying observations in the series, and Mandelbrot discovered that by using stable distributions with an α parameter of less than 2, the parameter of the normal distribution, these "fatter tailed" distributions fit better than the normal. Fama lent further support to the statistical case for the stable distributions and, as did Samuelson, developed some portfolio theoretic implications when returns were stable, but not necessarily normal. Fama also generalized the mean variance capital asset pricing model of Sharpe and Lintner to the case where returns were generated by a multivariate stable factor model.

For a variety of reasons, though, the stable distributions have not proved to be an altogether satisfactory generalization of current mean variance financial models. On an empirical basis, Clark has demonstrated that as a pure matter of goodness of fit the subordinated normal processes also generate the outlying observations in speculative price data and, in general, outperform the stable distributions in explaining the data. On theoretical grounds things are even worse.

The first difficulty noted with the stable distributions was that, since they do not possess moments of order greater than or equal to their α

parameter ($0 < \alpha < 2$), von-Neumann Morgenstern expected utilities would not exist for a number of conventional utility functions. Of course, one can take the view that this is so much the worse for those utility functions, but this perspective gets a bit cramped as α declines to unity and one runs out of utility functions for which expected utility exists (see Ross and Blume). Somewhat more significantly, though, it was originally thought that the stable distributions would, in some sense, be a unique generalization of the mean variance model of financial theory, and this has proven to be incorrect. Ross has completely delineated the class of distributions that permit such generalizations and it is significantly larger than the stable class.

Nevertheless, the stable distributions do retain a tantalizing appeal. For one thing they have canonical properties as the limit laws of sequences of independent random variables. The famous Khintchine-Levy theorem characterized the stable laws and they are easily shown to be the only ones which can be the limit distributions for sums of independently and identically distributed random variables (see Gnedenko and Kolmogorov). This limiting property is suggestive in a context where prices are determined by demand and supply conditions that are made up of a large number of individual decisions, but within the context of neoclassical portfolio theory there really is no longer any unique role for the stable distributions.

The intent of this paper is to suggest a theoretical basis for such a role. In particular, the stable distributions will be shown to be the unique ones which satisfy a small set of simple assumptions on portfolio behavior. In so doing, we will be able to link up two important concepts in financial theory, the stable distributions and portfolio separation--stability and separability for short. Section I introduces these concepts and Section II proves the central theorems linking them together. Section III considers extensions of these results with some implications for multivariate statistical theory. Section IV briefly concludes the paper with a discussion of unresolved issues.

I. Stability and Separability--Definitions

The purpose of this section is simply to introduce the two key notions that will be used in the paper. The first concept is that of stability and we will include some results from the theory of stable law solely for the purpose of keeping the paper self contained. (These results are taken from Breiman, Feller, and Gnedenko and Kolmogorov.)

A type of distribution function is a set of distribution functions generated by changes of scale and location. If $F(x)$ denotes a distribution function then the type generated by $F(x)$ is the class of distribution functions of the form $F(ax + b)$, where $a > 0$ and b are arbitrary constants. For example, if we change the scale and shift the location of a normal random variable, it remains normal; hence the normals form a type. A type is called stable if it is closed under composition, i.e., a type is stable if the sum of any two independent members of the type belongs to the type. Thus, if x_1 and x_2 are any two independent random variables in the same type, the type is stable only if

$$x_1 + x_2 = x_3$$

is also in that type. To put the same thing in the form of distribution functions, if $F(a_1x + b_1)$ and $F(a_2x + b_2)$ are the distribution functions of x_1 and x_2 respectively (a_1, $a_2 > 0$), then stability requires that there exist constants $a_3 > 0$ and b_3 such that $F(a_3x + b_3)$ is the distribution function of x_3. Since the sum of two independent normals is also normal, the normal type is a stable type.

In a famous theorem, Khintchine and Levy showed that a distribution function is stable if and only if the logarithm of its characteristic function[1] is of the form

$$\log f(w) = i\gamma w - c|w|^{\alpha}\{1 + i\beta\frac{w}{|w|}\,\omega(w,\alpha)\}, \qquad (1)$$

where α, β, γ and c are constants,

$$-1 \leq \beta \leq 1,$$

$$c \geq 0.$$

The α parameter which tells the highest absolute moment possessed by the distribution function (except for $\alpha = 2$ which corresponds to the normal distribution which possesses all moments) belongs to the interval [0, 2],

and

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

The constants in (1) have useful interpretations; γ is a measure of location the mean if the distribution has one (i.e., $a > 1$); α is a measure of spread; and β is a measure of skewness, the distribution being symmetric if $\beta = 0$. Unfortunately, though, while the characteristic function (1) is easily displayed in closed form and although a number of stable distribution functions besides the normal are known explicitly, there does not exist any simple characterization of the form of the distribution functions of the stable types.

The Khintchine and Levy theorem has a number of interesting consequences. One we shall use is the following. If $(x_1, \ldots, x_n)$ are identically and independently distributed ($\gamma = 0$) stable random variables with parameter α, then

$$x_1 + \cdots + x_n,$$

is distributed as a stable random variable ax_i with

$$a = n^{1/\alpha} . \qquad (2)$$

In the normal case this is easy to see since $x_1 + \cdots + x_n$ is normal with variance equal to $n\sigma^2$ (where σ^2 is the variance of each x_i), and this is simply the variance of $\sqrt{n}\, x_i$, and (2) may be verified in general from (1).

The second major concept we will use in this paper is that of separability in portfolio choice. It is well known, for example, that in a mean variance model with a riskless asset all investors with the same probability assessment of returns will divide their portfolios between the riskless asset and a common portfolio of risky securities. Attitudes towards risk enter, then, not to determine the composition of the risky portfolio, but, rather, only to determine the amount of wealth placed in the riskless asset and in the common risky portfolio. This sort of result is central to much of modern financial theory and is called a separation property. Ross (2) described those distributions which have the separation property and discussed a variety of types of separation.

The mean variance example illustrates two fund separability (2fs) in that all portfolios are divided between the common risky portfolio and the portfolio containing only the riskless asset. Even simpler is the concept of one fund separability (1fs) in which all risk averse expected utility maximizers choose the <u>same</u> one fund of assets. Suppose, for example, that there are only two assets with returns x_1 and x_2. From Ross (2) a necessary and sufficient condition for the portfolio $(\frac{1}{1+b}, \frac{b}{1+b})$ to be the optimal portfolio for all (risk averse) investors is that

$$E\{x_1 - x_2 | x_1 + bx_2\} = 0, \text{ (a.e.)}, \tag{3}$$

where E denotes the expectation operator. The necessity of this result follows from an argument in the stochastic dominance literature, but the sufficiency is easy to see. Suppose (3) holds. Let U be any monotone concave utility function and ξ any portfolio alternative to $(\frac{1}{1+b}, \frac{b}{1+b})$. Now,

$$(\xi_1, \xi_2) = (\frac{1}{1+b}, \frac{b}{1+b}) + (\eta_1, \eta_2)$$

defines (η_1, η_2) and $\eta_1 + \eta_2 = 0$ since $\xi_1 + \xi_2 = 1$. Hence

$$\xi_1 x_1 + \xi_2 x_2 = \frac{1}{1+b} x_1 + \frac{b}{1+b} x_2 + \eta_1 x_1 + \eta_2 x_2$$

$$= \frac{1}{1+b} x_1 + \frac{1}{1+b} x_2 + \eta_1 (x_1 - x_2).$$

From (3), then, $\xi_1 x_1 + \xi_2 x_2$ is distributed as

$$\frac{1}{1+b} x_1 + \frac{b}{1+b} x_2 + \text{(noise)},$$

and must therefore be inferior to the portfolio $(\frac{1}{1+b}, \frac{b}{1+b})$. Formally,

$$E\{U[\xi_1 x_1 + \xi_2 x_2]\} = E\{U[\frac{1}{1+b} x_1 + \frac{b}{1+b} x_2 + \eta_1 (x_1 - x_2)]\}$$

$$= E_{(\frac{1}{1+b} x_1 + \frac{b}{1+b} x_2)} \{E\{U[\frac{1}{1+b} x_1 + \frac{b}{1+b} x_2 + \eta_1 (x_1 - x_2)] \; \frac{1}{1+b} x_1 + \frac{b}{1+b} x_2\}\}$$

$$\leq E\{U[\frac{1}{1+b} x_1 + \frac{b}{1+b} x_2]\},$$

where we have used (3) and Jensen's inequality.

In the next section we will use the concept of one fund separability, as embodied in (3), to characterize the class of stable distributions.

II. Basic Results

The main result of this paper is the following theorem and its generalizations which link the concepts of separability and stability.

<u>Theorem 1:</u> (A Two Asset Theorem)

Let x_1 and x_2 be two independent random variables possessing means. A necessary and sufficient condition for the pair (x_1, ax_2) to exhibit lfs for all $a \geq 0$ is that x_1 and x_2 are distributed as mean zero stable random variables which can differ only in scale, i.e., they have common exponents $\alpha > 1$ and skew parameter β. The pair (x_1, ax_2) exhibits lfs for all values of a iff x_1 and x_2 are distributed as mean zero symmetric stable $(\beta = 0)$ random variables with common $\alpha > 1$.

In other words, with independent assets the stable distributions are the only ones which retain the separation property under changes of scale. The proof of this theorem is not simple and it requires a bit of work. The basic idea is to show that if x_1 and x_2 are stable with a common α, then (3) must hold and, conversely, that (3) implies that x_1 and x_2 are stable. The sufficiency part is generally thought to be well known, but it is instructive to work it out in a somewhat different fashion than usual. Notice that we require $\alpha > 1$ simply to permit expectations to exist. For a treatment of the esoteric case when $\alpha \leq 1$ see Ross and Blume.

The primary results we need to prove sufficiency are (2) and the following lemma which gives a set of conditions under which conditional expectations are zero.

Lemma:

Let $x_1,\ldots,x_n$ be independent, identically distributed random variables. For any constants $\eta_1,\ldots,\eta_n$,

$$\sum_{i=1}^{n} \eta_i = 0, \tag{4}$$

implies that

$$E\{\sum_{i=1}^{n} \eta_i x_i \mid \sum_{i=1}^{n} x_i\} = 0. \tag{5}$$

Proof:

Consider all permutations, k, of $(\eta_1,\ldots,\eta_n)$. For any such permutation

$$E\{\sum_{i=1}^{n} \eta_{k_i} x_i \mid \sum_{i=1}^{n} x_i\} = c,$$

independent of the permutation, k, since

$$\sum_{i=1}^{n} x_i,$$

can give no differential information about the individual x_i's. Letting # denote the total number of permutations and using (4) we have

$$\begin{aligned} \# \cdot c &= E\{\sum_k \sum_{i=1}^{n} \eta_{k_i} x_i \mid \sum_{i=1}^{n} x_i\} \\ &= E\{\sum_{i=1}^{n} (\sum_k \eta_{k_i}) x_i \mid \sum_{i=1}^{n} x_i\} \\ &= E\{\sum_{i=1}^{n} (0) x_i \mid \sum_{i=1}^{n} x_i\} \\ &= 0, \end{aligned}$$

hence

$$c = 0.$$

Q.E.D.

Armed with this lemma we will prove the sufficiency part of Theorem 1.

Proof of Theorem 1:

Sufficiency:

From (3) it suffices to show that for any $a > 0$ there exists a b ($\neq -1$) such that

$$E\{x_1 - ax_2 | x_1 + bx_2\} = 0,$$

the symmetric case follows since x_j and $-x_j$ are identically distributed.

Suppose first that the quantity $a^{\frac{\alpha}{\alpha-1}}$ is rational, and let m and n be integers with

$$a^{\frac{\alpha}{\alpha-1}} = \frac{m}{n},$$

($a^{\frac{\alpha}{\alpha-1}}$ is well defined since $\alpha > 1$). Choose b to be

$$b = \left(\frac{n}{m}\right)^{\frac{1}{\alpha}} .$$

Using (2) there exist m+n identically distributed, independent, stable random variables, $y_1,\ldots,y_m$ and $z_1,\ldots,z_n$ such that

$$x_1 = y_1 + \cdots + y_m,$$

and (6)

$$bx_2 = z_1 + \cdots + z_n,$$

where each y_i and z_i is distributed as

$$\frac{1}{m^{1/\alpha}} x_1$$

(or $m^{1/\alpha}$ x_2 since x_1 and x_2 are identically distributed).

Hence,

$$x_1 - ax_2 = y_1 + \ldots + y_m - \frac{a}{b}(z_1 + \ldots + z_n),$$

and since

$$n\left(\frac{a}{b}\right) = \left(\frac{m}{n}\right)^{\frac{\alpha-1}{\alpha}} \cdot \left(\frac{m}{n}\right)^{\frac{1}{\alpha}} \cdot n = m,$$

we can apply the lemma to show that

$$E\{x_1 - ax_2 | x_1 + bx_2\} = 0.$$

Having proved the result for all rational $a^{\frac{\alpha}{\alpha-1}}$, the theorem follows by a straightforward closure argument.

Necessity:

The characteristic functions of x_1 and x_2 are defined by

$$f_1(w) = E\{e^{wx_1}\}$$

and

$$f_2(w) = E\{e^{wx_2}\} .$$

If lfs holds for a pair (a, b) then by (3)

$$E\{x_1 - ax_2 | x_1 + bx_2\} = 0,$$

and, as a consequence, x_1 and x_2 are mean zero, and it also follows that for all w,

$$E\{e^{iw(x_1+bx_2)}(x_1 - ax_2)\} = 0. \tag{7}$$

Now, since x_1 and x_2 are independent,

$$E\{e^{iw(x_1+bx_2)}x_1\} = E\{e^{iwx_1}x_1\}E\{e^{iwbx_2}\}$$

$$= -i\ f_1'(w)f_2(wb),$$

and

$$E\{e^{iw(x_1+bx_2)}x_2\} = E\{e^{iwx_1}\}E\{e^{iwbx_2}x_2\}$$

$$= -i\ f_1(w)f_2'(wb).$$

Hence, from (7),

$$f_1'(w)f_2(wb) = af_1(w)f_2'(wb),$$

or

$$\frac{f_1'(w)}{f_1(w)} = a\frac{f_2'(wb)}{f_2(wb)}, \tag{8}$$

where $f_j(z) \neq 0$. The assumption of a first moment guarantees that the first derivative of the characteristic function exists and is continous. Since f_j is a characteristic function, $f_j(0) = 1$, and, therefore, $f_j \neq 0$ on some open neighborhood, I, of the origin. Defining

$$m_j(w) \equiv \frac{f_j'(w)}{f_j(w)}, \tag{9}$$

on I, we can now rewrite (8) as the functional equation

$$m_1(w) = am_2(wb). \tag{10}$$

Suppose, first, that a is unrestricted. In the appendix we show that the only admissible solutions to this equation are

$$m_j(w) = c_j' \frac{w}{|w|} |w|^d, \tag{11}$$

where c_j' and d are constants.

From (8) and (11), then, integration yields

$$\log f_j(w) = c_j' \frac{1}{d+1} |w|^{d+1}, \tag{12}$$

where the requirement that $f_j(0) = 1$ eliminates the constant of integration and assures that $d \neq -1$. If $c_1 = c_2 = 0$, then $f_j(w) = 1$ and x_1 and x_2 are identically zero, which is the degenerate stable case.

Defining $\alpha \equiv d + 1$ and $c_j = -\frac{1}{\alpha} c_j'$, (12) may be rewritten as

$$\log f_j(w) = -c_j |w|^\alpha. \tag{13}$$

Now, (13) is a local result which holds on I. Let $w^* = \sup\{|w| (13) \text{ holds}\}$. If $w^* < \infty$, then $f_j(w) = 0$ for $|w| > |w^*|$ and arbitrarily close to w^*. But with (13) this would violate the continuity of f_j at $|w^*|$, and (13) must hold globally. It only remains to show that (13) defines a symmetric stable random variable. Since x_j has an expectation, we must have $\alpha > 1$, and since $\alpha > 2$ would imply, by (13), a zero second moment, we must also have $\alpha \leq 2$.

Since $|w|^{\alpha}$ is an even function, c_j is real and $c_j > 0$ is required to bound f_j in modulus ($c_j = 0$ is the trivia distribution with unit mass at zero). This completes the description of a symmetric stable characteristic function.

Suppose, now, that we restrict $a \geq 0$. From the appendix the solutions to (10) are given by

$$m_j(w) = c_j^{\pm} |w|^d, \text{ as } w \gtrless 0,$$

where (14)

$$\frac{c_1^+}{c_2^+} = \frac{c_1^-}{c_2^-} \quad .$$

Integrating (14) yields

$$\log f_j(w) = \begin{cases} c_j^+ \frac{1}{d+1} |w|^{d+1} & w \geq 0, \\ -c_j^- \frac{1}{d+1} |w|^{d+1} & w < 0. \end{cases} \tag{15}$$

Again an extension argument verifies that (15) holds globally. Since the real and complex parts of a characteristic function are, respectively, even and odd functions, $-c_j^-$ is the complex conjugate of c_j^+ and we can rewrite (15) as

$$\log f_j(w) = -c_j |w|^{\alpha}\{1 + i\beta \frac{w}{|w|} \omega\} \tag{16}$$

where

$$c_j^+ \equiv -c_j\{1 + i\beta\omega\},$$

and

$$-c_j^- \equiv -c_j\{1 - i\beta\omega\},$$

and

$$\omega \equiv \tan \frac{\Pi}{2} \alpha \ .$$

From (14) β is independent of j and, as before, $c_j \geq 0$ and $1 < \alpha \leq 2$. Hence (16) coincides with the stable characteristic function (1), and x_1 and x_2 differ only in scale.[2]

Q.E.D.

Theorem 1 is readily generalized to the case of $n > 2$ random variables. One such generalization is given below.

Theorem 2: (An n- Asset Theorem)

Let x be a vector of n independent random variables possessing means and let a and b be n-vectors. The following conditions are equivalent.

(C1) For all choices of a, the subvector of

$$a \circ x \equiv (a_1 x_1, \ldots, a_n x_n),$$

exhibits ifs

and

(C2) The x_i are symmetric stable with zero mean and common characteristic exponent $\alpha > 1$.

The following conditions are also equivalent.

(C1)' For all choices of $a \geq 0$, the subvector of $a \circ x$ exhibits ifs

and

(C2)' The x_i are stable with zero mean and can differ only in scale.

Proof:

Assume that (C1) ((C1)') holds. Pick a (a $\geq$ 0). From ifs ($\exists$b) such that ($\forall$i, j)

$$E\{a_i x_i - a_j x_j \mid b(a \circ x)\} = 0,$$

or

$$E\{(a_i x_i - a_j x_j)\, e^{ib(a \circ x)}\} = 0,$$

but this implies that (8) must be satisfied with $a \equiv {}^{a_j}/a_i$ and $b \equiv \frac{b_j a_j}{b_i a_i}$, hence, by Theorem 1 (C2) ((C2)') must follow.

Conversely, if (C2) ((C2)') holds, then applying the lemma it is easily verified that the separating portfolio, b, will be given by

$$b_j = \lambda \frac{a_j}{|a_j|} \; |a_j|^{1/1-\alpha}, \tag{17}$$

where $\lambda > 0$ is an arbitrary constant.

Q.E.D.

Theorem 2 describes the unique position of the stable random variables in both multivariate probability theory and in portfolio theory. Roughly summarizing, a set of independent random variables will be separable at all levels of scale if and only if the variables are stable with the same characteristic exponent. The next section will examine some further implications of Theorem 2 and will consider some generalizations.

III. Some Extensions

A. Orthogonality

The multivariate normal distribution has proven to be a useful basis for portfolio theory largely because it possesses the mutual fund separation property (see Ross (2)), but, there are a number of other canonical properties possessed by both the normal and the stable distributions and it would be useful to see how these are related to separation.

For example, consider two normal random variables x_1 and x_2 with the same mean and a covariance matrix, V. If $b = (b_1, b_2)$ is any portfolio of x_1 and x_2, then any other portfolio a with

$$a\ V\ b = 0 \tag{18}$$

has a return, ax, uncorrelated with bx and, furthermore, its return is actually independent of the portfolio return bx. This property, that lack of correlation implies independence, provides an easy way to find the optimal portfolio, b. Simply let

$$b = \lambda V^{-1} e, \tag{19}$$

where λ is a scaling constant set to let $be = 1$, and e is a vector of ones. Now, any other portfolio a will have

$$\eta \equiv a - b\ , \tag{20}$$

where

$$\eta e = ae - be = 0, \tag{21}$$

and, by construction,

$$\eta\, V\, b = \lambda \eta e = 0. \qquad (22)$$

From the stochastic dominance rule, (3), then b must be the optimal portfolio since

$$E\{\eta x \mid bx\} = 0. \qquad (23)$$

The property of zero correlation is an orthogonality property, i.e., (18) is equivalent to rotating each vector by a transform $V^{1/2}$ (which exists since V is non-negative definite) and requiring orthogonality of the rotated vectors. For multivariate normal random variables the orthogonality condition, (18), implies that

$$E\{ax \mid bx\} = E\{ax\}. \qquad (24)$$

To make the point in a constructive fashion (18) assures us that given any a we can find a b such that (24) holds and, conversely, given a b we can find a such that (24) holds. These, seemingly weaker properties, are stated below in a formal fashion suitable for our cases.

(C3) (Right Orthogonality) We will say that x has the right orthogonality property (RO) if $(\forall a)$ $(\exists b$ with $a_i = 0 \Rightarrow b_i = 0$, for a with at least two nonzero elements)

$$E\{ax \mid bx\} = 0.$$

(C4) (Left Orthogonality) We will say that x has the left orthogonality propert (LO) if $(\forall b)$ ($\exists a$ with $b_i = 0 \Rightarrow a_i = 0$, for b with at least two nonzero elements.)

$$E\{ax \mid bx\} = 0.$$

Notice that the x are implicitly assumed to possess means in (C3) and (C4).

Despite the seeming weakness of (C3) and (C4) we can prove the following rather surprising conditions.

Theorem 3:

Assuming that x is a vector of independent random variables possessing means, conditions (C1) through (C4) are equivalent.

Proof:

From Theorem 2 (C1) and (C2) are equivalent. Assume that (C1) holds. For any choice of a scale vector, c, then, there exists b (with $c_i = 0 \Rightarrow b_i = 0$) such that for all η with $\eta e = 0$, (and $\eta_i = 0$ if $c_i = 0$)

$$E\{\eta(c \circ x) \mid b\,(c \circ x)\} = 0. \tag{25}$$

Now, for any a we can define c and η such that

$$a_i = \eta_i\, c_i\ , \tag{26}$$

hence, (25) implies (C3).

Condition (C4) follows directly (C2) and (17).

Consider any b vector and let

$$a_j = \lambda |b_j|^{\frac{1-\alpha}{2-\alpha}} . \tag{27}$$

From (17), for all η with $\eta e = 0$, (27) implies that

$$E\{\eta(a \circ x) \mid bx\} = 0 , \tag{28}$$

which verifies (C4).

Now, assume that (C3) holds. Consider a vector a with

$$a_i, a_j \neq 0$$

$$a_k = 0, \quad k \neq i, j.$$

Condition (C3) now takes the same form as in Theorem 1 and it follows that x_i and x_j satisfy (C2). Since this holds for all (x_i, x_j) pairs (C2) is satisfied.

Similarly if (C4) holds, then we can also apply Theorem 1 to (x_i, x_j) pairs to verify (C2), since the direction of causality from a to b or b to a was irrelevant in the proof of Theorem 1.

<u>Q.E.D.</u>

There is also an asymetric version of conditions (C3) and (C4) and of Theorem 3 and we include it below.

(C3)' We will say that x has the (RO)' property if $(\forall a,\ ae = 0)$ $(\exists b$ with $a_i = 0 \Rightarrow b_i = 0$ for a with at least two nonzero elements).

$$E\{ax \mid bx\} = 0 \ . \tag{29}$$

(C4)' We will say that x has the (LO)' property if $(\forall b)$ $(\exists a,\ ae = 0,$ and $b_i = 0 \Rightarrow a_i = 0$ for b with at least two nonzero elements).

$$E\{ax \mid bx\} = 0. \tag{30}$$

Theorem 3':

Assuming that x is a vector of independent random variables possessing means, conditions (C1)' through (C4)' are equivalent.

Proof:

The proof is a straightforward adaptation of the proof of Theorem 3.

Q.E.D.

B. Dependence.

The complete extension of the above results to cases where x is a general random vector is a difficult task, but it is possible to treat an important subclass of cases that permits dependence among the random return.

One traditional way to build multivariate distributions with dependence is to construct them as sums of independent random variables. Let L denote the class of x for which $\exists$(A, u) such that

$$x = Au, \tag{31}$$

where A is a given matrix of full rank and u is a vector of independent random variables possessing means. It can be shown than not all multivariate random variables belong to L, but L is an important subclass of the set of multivariate random variables. It should also be clear that there is no loss of generality in making A of full rank.[3]

The next theorem extends the previous results to the class L.

Theorem 4:

Let $x \in L$. The following conditions are equivalent.

(C1)* The vector Sx exhibits ifs for all S generated by

$$S = CA^{-1},$$

where C is any diagonal matrix.

(C2)* The u_i are symmetric stable with mean zero and common characteristic exponent $\alpha > 1$.

$(C3)^*$ For all a, there exists b with $(aA)_i = 0 \Rightarrow (bA)_i = 0$ for (aA) with at least two nonzero elements, such that

$$E\{ax \mid bx\} = 0.$$

and

$(C4)^*$ For all b, there exists a with $(bA)_i = 0 \Rightarrow (aA)_i = 0$, for (bA) with at least two nonzero elements, such that

$$E\{ax \mid bx\} = 0.$$

<u>Proof:</u>

The proof is a straightforward exercise in cancelling out the A transform by inversion. To illustrate, assume $(C1)^*$. From (25) for all η, $\eta e = 0$

$$\begin{aligned} E\{\eta Sx \mid bSx\} &= E\{\eta CA^{-1} Au \mid b\, CA^{-1} Au\} \\ &= E\{\eta Cu \mid bCu\} \\ &= 0, \end{aligned}$$

which is a restatement of ifs for u, and, therefore implies $(C2)^*$.

<u>Q.E.D.</u>

Theorem 4':

Let $x \in L$. The conditions $(C1)^*$ through $(C4)^*$ are equivalent when modified so that

(i) $C \geq 0$,

(ii) $(C2)'$ replaces $(C2)^*$,

and

(iii) $aAe = 0$.

Proof:

A simple modification of the proofs of Theorem 3' and Theorem 4.

Theorem 4 and 4' can actually be strengthened and, in particular $(C2)^*$ implies that the restrictions on S in $(C1)^*$ can be eliminated. We will prove this in the next section in the context of a somewhat different set of results.

C. Regression and Capital Asset Pricing Models.

The two parameter risk and return capital asset pricing model (CAPM) has become the workhorse for both theoretical and empirical exercises on capital markets. The conclusions of such models are usually summarized in an equilibrium pricing relation of the form

$$E_i = E_o + \lambda b_i \,, \tag{32}$$

where λ is a constant, E_o is the risk free return, E_i is the expected return on the i^{th} asset and b_i is a measure of the covariation of the return x_i, with some reference portfolio such as the market portfolio of all risky assets. (For derivations of (32) see Sharpe, Lintner, Black or Ross (1)).

A traditional first step in empirically testing the model is to obtain estimates of E and b by fitting the following (time series) regression to the data, x:

$$x = E + b(ax) + \varepsilon, \tag{33}$$

where a is the market portfolio and

$$E\{\varepsilon | ax\} = 0. \tag{34}$$

The estimated coefficients b are the beta coefficients used in testing the CAPM. If x is multivariate normal, then for arbitrary a if b is the vector of covariances of x with ax, (34) will hold. In other words, if (33) defines ε, ε will satisfy (34). On the other hand, for arbitrary distributions of x there will not generally exist a vector b such that using (33) to define ε, (34) will hold. In at least this sense, then, multivariate normality is a sufficient condition for the regression to be valid. This property of having residuals which have zero conditional mean, (34), is also of importance in factor analysis where factors are contructed as linear combinations of the random observations. To formalize this concept we will define the following linear regression property.

(C5)* A multivariate random vector, x, possessing means, is said to have the linear regression property (LRP) if and only if $(\forall a)$ $(\exists b)$ such that if (33) defines ε, then (34) holds, i.e., for each i

$$E\{x_i | ax\} = b_i ax. \tag{35}$$

We can now use our previous results to describe an important subclass of the class of multivariate distributions with LRP.

Theorem 5:

If $x \in L$, then x has the LRP if and only if the u in (31) are mean zero symmetric stable with common characteristic exponent $\alpha > 1$, i.e., conditions $(C1)^*$ through $(C5)^*$ are equivalent.

Proof:

Consider sufficiency first. From Theorem 4, $(C1)^*$ through $(C4)^*$ are equivalent, and using $(C2)^*$, let the u_i be symmetric stable and let a be any vector. Equivalent to (35) we have to show that $(\forall a)\ (\exists b)\ cb = 0$ implies

$$E\{cx|ax\} = 0, \qquad (36)$$

or in terms of u we want

$$E\{cAu|aAu\} = 0. \qquad (37)$$

In the u's differ in scale we will include this scale transform in A, hence, without loss of generality, let the u_i be identically distributed. This permits us to rewrite the LRP as requiring that $(\forall a)\ (\exists b)$, $cb = 0$ implies

$$E\{cu|au\} = 0. \qquad (38)$$

Since the u_i are symmetric stable, we can convert au into a sum of independent identical stable random variables by the same procedure we used to verify sufficiency in Theorem 1. The lemma now provides the b vector for which (38) holds. Proceeding formally, suppose first that the $|a_i|^\alpha$ are rational. If follows that ($\exists m_i$ and m)

$$|a|^{\alpha} = \frac{m_i}{m} ,$$

and by (2) and symmetry au is distributed as

$$\sum_{i=1}^{n} |a_i| u_i = \sum_{i=1}^{n} \{ \sum_{j=1}^{m_i} z_{ij} \} ,$$

where the z_{ij} are independent and distributed as $(\frac{1}{m})^{1/\alpha} u_i$. By Lemma 1, then

$$E\{cu|au\} = E\{ \sum_{i=1}^{n} y_i \frac{a_i}{|a_i|} (\sum_{j=1}^{m_i} z_{ij}) | \sum_{i,j} z_{ij} \}$$

$$= 0,$$

if

$$\sum_{i=1}^{n} m_i \frac{a_i}{|a_i|} y_i = 0, \tag{39}$$

where

$$y_i \equiv \begin{cases} \frac{c_i}{|a_i|} & \text{if } |a_i| \neq 0, \\ 0 \text{ if } |a_i| = 0. \end{cases}$$

We can rewrite (39) as

$$0 = \sum_{i=1}^{n} m_i y_i = \frac{1}{n} \sum_{i=1}^{n} m_i \frac{a_i}{|a_i|} y_i$$

$$= \sum_{i=1}^{n} |a_i|^{\alpha-1} \frac{a_i}{|a_i|} c_i \tag{40}$$

$$= 0,$$

and since the rationals are dense, the vector

$$b_a = \left(\frac{a_1}{|a_1|} \cdot |a_1|^{\alpha-1}, \ldots, \frac{a_n}{|a_n|} |a_n|^{\alpha-1}\right) \qquad (41)$$

is the one we require.

Now consider the necessity portion of the proof.

Let a_i , $a_j \neq 0$, $a_k = 0$, $k \neq i, j$. Now, (38) becomes

$$E\{c_i u_i + c_j u_j \mid a_i u_i + a_j u_j\} = 0, \qquad (42)$$

where

$$c_i b_i + c_j b_j + \sum_{k \neq i,j} c_k b_k = 0,$$

and setting $c_k = 0$, $k \neq i,j$ we have

$$c_i b_i + c_j b_j = 0. \qquad (43)$$

Since either c_i or c_j can be taken to be nonzero, condition (42) is equivalent to (C4) on pairs. From Theorem 3 (C2)* must hold on pairs, and, as a consequence, all the u_i must be symmetric stable with a common characteristic exponent $\alpha > 1$. Theorem 4 completes the proof.

Q.E.D.

The asymmetric version of the LRP, (C5)', restricts $aA \geq 0$. The following theorem provides the asymmetric form of Theorem 5.

Theorem 5':

If $x \in L$, then x has (C5) if and only if the u in (31) are stable and can differ only in scale, i.e., the primed versions of $(C1)^*$ through $(C5)^*$ are equivalent.

Proof:

See the proofs of Theorems 4 and 5.

Q.E.D.

In an important sense, then, Theorems 5 and 5' severely limit the class of multivariate distributions for which factor analytic regressions of the form of (33) are possible. This, in turn, has strong implications for the multivariate distributions implicitly assumed in empirical tests of the CAPM. Basically, when regressions are run on arbitrary portfolios, the assumption of multivariate stability is implicit.

Finally, we should remark on the relationship of the results of this section to the more traditional approach of defining the multivariate stable distributions. Levy and Feldheim gave an explicit characterization of this form, for vector x, analogous to (1) and defined by the property of stability of the law under addition, scale and location changes. Press has developed these results further and proven some useful properties of the multivariate stable defined in this fashion. For example, Press has shown that a random vector x follows a Levy multivariate stable distribution if and only if all linear combinations of the x_i are univariate stable. It follows, then, that if the u_i in (31) are stable, then x is a Levy multivariate stable random variable. I do not know whether or not the converse is true. As we have seen, though, multivariate stable random variables as defined by L have a number of important properties, and if this class is a proper subset of the Levy class then it remains to be determined whether the wider class of multivariate stable random variables also retains these properties.

IV. Conclusions

This paper has examined some links between the concepts of stability in probability theory and separability in portfolio theory. Ross (2) showed that portfolio separation did not imply distributional stability, but this paper has shown that if a random vector of asset returns is to be separable (1fs) at all scales of operation, then it must be stable. Section III developed some implications of these results for regression and factor analysis, and multivariate statistical theory, and applied these results to the problems of testing capital asset models. In particular, we showed that a multivariate distribution must be stable for regression analysis to be applied on all possible portfolios. There remain, however, a number of unresolved questions.

To mention two of the most important issues, first, the extension of these findings to the (2fs) case (and more generally, the (kfs) case) should lead to important results. As a conjecture, if the above results were extended to include location and scale changes, the stable distributions would again play a necessary role and would be permitted to have arbitrary means. Secondly, while the results have been extended to an important subclass of multivariate random variables with dependence permitted, a full understanding of the general multivariate case is not yet available. Further work along these lines should broaden our knowledge of the interaction between portfolio theory and statistical theory.

Appendix

This appendix provides a solution to the functional equation (10) that arises in the proof of Theorem 1. The proof is adapted from Breiman and apparently owes its origin to Cauchy. We should note that one direct way of solving equations like (10) is to differentiate them in an effort to turn them into differential equations. Such an approach would give the same solution that we find, but is not really proper since the derivative of m_j involves the second derivative of f_j which does not exist for stable distributions with $\alpha < 2$ and which we cannot a priori assume exists. For simplicity we change notation a bit in Theorem A1 below.

Theorem A1:

Suppose that there is an open interval, I, containing the origin on which $(\forall a)(\exists b)$ such that

$$f(w) = ag(bw), \tag{a1}$$

where f and g are continuous. It follows that there exist constants c, k and d such that

$$\begin{aligned} f(w) &= c\frac{w}{|w|}\,|w|^{d}, \\ g(w) &= k\frac{w}{|w|}\,|w|^{d}, \end{aligned} \tag{a2}$$

and

$$b_a = \frac{c/ka}{|c/ka|}\,|c/ka|^{1/d} \quad .$$

Proof:

We first convert (al) into a functional equation in f alone. If $b = 0$ for all a ($\neq 0$) then $f = g = 0$. Suppose, then, that ($z \in I$) with $b_z \neq 0$. From (al)

$$f(w) = ag(wb_a) = zg(wb_z),$$

or

$$xf(w) = f(wc_x), \qquad \text{(a3)}$$

where we have defined

$$x \equiv {}^{z}/_{a},$$

and

$$c_x \equiv {}^{b_a}/_{b_z} .$$

Since for any interval of x, w can be chosen sufficiently small so that w, $wc_x \in I$, c_x is a global relation and can be solved for without regard to I. From (a3) and continuity c_x is monotone, hence by (a3)

$$f(wc_xc) = yf(wc_x) = xyf(w)$$

implies that

$$c_{xy} = c_xc_y . \qquad \text{(a4)}$$

Now, (a4) is the familiar Canchy equation and, for completeness alone, we will work out its solution. First, for any integer n, using (a4)

$$c_2 \equiv c(2) = c([2^{1/n}]^n)$$

$$= [c(2^{1/n})]^n,$$

hence

$$c(2^{1/n}) = [c_2]^{1/n} .$$

It follows that for any integers m, n

$$c(2^{\frac{m}{n}}) = c((2^{\frac{1}{n}})^m)$$

$$= c(2^{\frac{1}{n}})]^m$$

$$= [c_2]^{\frac{m}{n}} \quad .$$

Since $2^{\frac{m}{n}}$ is dense on the nonnegative line, for all $x > 0$

$$c(x) = [c_2]^{\log_2 x} ,$$

or

$$\log_2 c(x) = \log_2 x \log_2 c_2$$

$$= \log_2 x^{\log_2 c_2} ,$$

which implies that

$$c(x) = x^e$$

where (a5)

$$e \equiv \log_2 c_2 \ .$$

(Notice that $x > 0$ implies $c_x > 0$ from (a3).)

Now consider c_{-x}. From (a3)

$$c_{-1} = c_{-1 \cdot 1} = c_{-1} \cdot c_1$$

or

$$c_1 = 1,$$

and

$$c_1 = 1 = c_{-1\cdot -1} = c_{-1} \cdot c_{-1}$$

implies

$$c_{-1} = \pm 1.$$

Thus,

$$c_{-x} = c_{-1\cdot x} = c_{-1} \cdot c_x = \pm c_x ,$$

and to preserve monotonicity we must have

$$c_{-x} = -c_x. \tag{a6}$$

Hence, for $x \varepsilon R$

$$c_x = \frac{x}{|x|} |x|^e. \tag{a7}$$

From (a3), if we hold

$$wc_x = y$$

a constant, then

$$f(w) = \frac{f(y)}{x}$$

$$= \frac{y/w}{|y/w|} |y|^{-1/e} f(y)|w|^{1/e}$$

$$\equiv c\frac{w}{|w|} |w|^d. \tag{a8}$$

Similarly, from (a1)

$$g(w) = k\frac{w}{|w|}\,|w|^{d}, \tag{a8}$$

and

$$b = \frac{c/ka}{|c/ka|}\,|c/ka|^{1/d} \quad . \tag{a8}$$

Q.E.D.

If we restrict a to be one sided, i.e., $a \geq 0$, then a nearly identical argument to the one given above verifies the following result.

Theorem A2:

Suppose that there is an open interval, I, containing the origin, on which $(\forall a > 0)(\exists b)$ such that (a1) holds. It follows that there exist constants $c_{\pm}$, $k_{\pm}$ and d such that

$$f(w) = \begin{cases} c_{+}w^{d} \text{ if } w \geq 0, \\ c_{-}|w|^{d} \text{ if } w < 0, \end{cases}$$

$$g(w) = \begin{cases} k_{+}w^{d} \text{ if } w \geq 0, \\ k_{-}|w|^{d} \text{ if } w < 0, \end{cases} \tag{a9}$$

and

$$b_{a} = \frac{c/ka}{|c/ka|}\,|c/ka|^{1/d} \quad ,$$

where

$$\frac{c}{k} \equiv \frac{c+}{k+} = \frac{c-}{k-} \quad .$$

Proof:

The argument of Theorem A1 leading to (a5) is unaltered, where we take $y > 0$ and $x \equiv {}^{y}/_{a} > 0$ for $a > 0$. The argument leading to (a8), however, is changed by the need to specify different constants y_- and y_+ such that

$$wc_x = y_- ,$$

if $w < 0$, and

$$wc_x = y_+ ,$$

if $w > 0$, since c_x is one signed (positive). This gives us the freedom to specify pairs of constants $c_{\pm}$ and $k_{\pm}$ depending on whether w is positive or negative. The requirement that

$$\frac{c_+}{k_+} = \frac{c_-}{k_-} ,$$

is necessary to define b_a consistently.

Q.E.D.

FOOTNOTES

1. The characteristic function, $f(x)$, of a random variable z is defined by $f(w) \equiv E\{e^{iwx}\}$.

2. I have not verified directly that $-1 \leq \beta \leq 1$ as required in (1), and I am somewhat puzzled by the force of this restriction. As Feller remarks on p. 542, the bounding of $|\beta|$ is "the surprising feature of the theorem." However, the restriction would seem to be a requirement for (1) to be the logarithm of a characteristic function (i.e., a positive definite function). If not, that is, if for some $|\beta| > 1$ (1) were still the logarithm of an acceptable characteristic function, f, then it is easily verified that f would be stable invalidating the Khintchine and Levy theorem. The same reasoning applies to (16) and verifies that $|\beta| \leq 1$.

3. If A were not of full column rank then we could combine u_i to reduce the rank and if there were more rows than columns we can simply add random variables that are 0 a.e.

REFERENCES

Black, Fischer, "Capital Market Equilibrium with Restricted Borrowing," *Journal of Business*, March 1973.

Breiman, Leo, *Probability* Massachusetts: Addison-Wesley Publishing Company, 1968.

Clark, Peter, "A Subordinated Stochastic Process Model with Finite Variance for Speculative Markets," *Econometrica* January 1973, 147-157.

Fama, E., "Risk, Return and Equilibrium: Some Clarifying Comments," *Journal of Finance*, 1968, 29-40.

Feldheim, M.E., "Étude de la Stabilité des Lois Probabilité," (Thése de la Faculté des Sciences de Paris), 1937.

Feller, William, *An Introduction to Probability Theory and its Applications Vol. II* New York: John Wiley & Sons, Inc., 1966.

Gnedenko, B.V. and A.N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables* Massachusetts: Addison-Wesley Publishing Company, 1954.

Levy, P., *Theorie de l'Addition des Variables Aléatoires* Paris: Gauthier-Villers; Second Edition, 1954.

Lintner, John, "The Valuation of Risky Assets and the Selection of Risky Investment in Stock Portfolios and Capital Budgets," *Review of Economics and Statistics*, 1965.

Mandelbrot, B., "The Variation of Certain Speculative Prices," *Journal of Business*, 36, 1963, 394-411.

Press, S.J., "Multivariate Stable Distributions," Center for Mathematical Studies in Business and Economics, University of Chicago, Report No. 7228, June 1972.

Ross, Stephen (1), "The Aribtrage Theory of Capital Asset Pricing," *Journal of Economic Theory*, forthcoming 1976.

Ross, Stephen (2), "Mutual Fund Separation in Financial Theory-- The Separating Distributions," Rodney L. White Center Working Paper, University of Pennsylvania.

Ross, Stephen and M. Blume, "A Note on a Paradox in Portfolio Theory"

Samuelson, Paul A., "Efficient Portfolio Selection for Pareto-Levy Investments," *Journal of Financial and Quantitative Analysis*, June 1967.

Sharpe, W., "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *Journal of Finance*, 19, 1964, 425-442.