The Annals of Statistics
1981, Vol. 9, No. 2, 391–400

ON THE PROJECTIONS OF ISOTROPIC DISTRIBUTIONS[1]

BY MORRIS L. EATON

University of Minnesota

The class of distributions on R^m, $1 \le m < \infty$ which are the m-dimensional marginal distributions of orthogonally invariant distributions on R^{m+n} is characterized. This result is then used to provide a partial answer to the following question: given a symmetric distribution on R^1 and an integer $n \ge 2$, under what conditions will there exist a random vector $X \in R^n$ such that $a'X$ has the given distribution (up to a positive scale factor) for all $a \ne 0$, $a \in R^m$.

1. Introduction. The present version of this paper resulted from a set of unusual circumstances. While teaching a course in distribution theory, the following question arose:

Q1. Given a symmetric distribution on the line and an integer $n \ge 2$, when will there exist a random vector X in R^n such that $a'X = \sum_1^n a_i X_i$ has the given distribution (up to a strictly positive scale factor) on R^1 for each vector $a \in R^n$, $a \ne 0$?

An answer to this question would allow one to define "n-dimensional versions" (as is done in the normal case) of symmetric distributions on R^1. This and some related questions were partially answered in Eaton (1977) and the results were submitted for publication. Shortly thereafter, Richard Olshen informed me that he had a set of handwritten notes of Leonard J. Savage (1969) concerning "round" distributions (orthogonally invariant or isotropic distributions) on R^n and there was some overlap with my work.

After going through the Savage notes, it seemed appropriate to prepare a paper which combined the work in Eaton (1977) and the results of Savage most directly related to Q1. Round distributions were the main topic treated in Savage's notes and one of the motivating questions was:

Q2. Which distributions on R^m can arise as the m-dimensional marginal distribution of an orthogonally invariant distribution on R^{m+n}?

In Section 2, we set some notation and describe a few "well-known" results to be used in the sequel. In Section 3, the basic representation result (Theorem 1) is established on R^m. This result was proved in Eaton (1977) for R^1 in attempting to answer Q1 and was proved by Savage (1969) for R^m enroute to his results in Section 4. The proof of Theorem 1 is a minor modification of the proof for R^1 given in Eaton (1977). Again, Theorem 2 in Section 3 was proved for R^1 in Eaton (1977) and was established in Savage (1969) for R^m. It seems clear that the methods developed in Freedman (1963) (see page 1194) can also be used to prove Theorem 2. The proof given here, which proceeds largely from first principles, is modelled after that for R^1 in Eaton (1977).

The results of Section 4, taken mainly from Savage (1969), give an analytic answer to Q2. Savage was not aware of Williamson's (1956) work and presented independent arguments which lead to Propositions 1 and 2. However, some errors in Savage's calculations necessitated that Propositions 1 and 2 be stated differently than the corresponding assertions in Savage's notes.

Received July 1977; revised November 1979.

[1] This work was supported in part by National Science Foundation Grant MCS77-25112.

AMS 1970 *subject classifications.* Primary 60E05, 62E10; secondary 62H05.

Key words and phrases. n-dimensional distributions, scale mixtures of normals, orthogonally invariant distributions.

The material of Section 5 is taken from Eaton (1977). The main result of this section (Theorem 3) gives a complete description of all the n-dimensional versions of a finite variance symmetric distribution on R^1. The case of infinite variance is discussed briefly and we close with a few remarks in Section 6.

2. Notation and background. In what follows, R^n denotes n-dimensional coordinate space and $\mathcal{O}_n$ is the group of $n \times n$ orthogonal matrices acting on R^n. Further let $S_n = \{x \mid \|x\| = 1, x \in R^n\}$ and let $B_n = \{x \mid \|x\| \leq 1, x \in R^n\}$ where $\|\cdot\|$ is the usual Euclidean norm on R^n. Of course, the following discussion can be given for any positive definite inner-product and norm on R^n. The distribution of a random vector $X \in R^n$ is denoted by $\mathscr{L}(X)$.

DEFINITION 1. For $X \in R^n$, the distribution $\mathscr{L}(X)$ is called *isotropic* if $\mathscr{L}(\Gamma X) = \mathscr{L}(X)$ for all $\Gamma \in \mathcal{O}_n$ and if $P\{X = 0\} = 0$. The class of all isotropic distributions on R^n is D_n.

The assumption that $P\{X = 0\} = 0$ in Definition 1 is simply to eliminate some uninteresting technical complications in the material that follows. Other names for isotropic distributions are orthogonally invariant distributions and Savage's term—round distributions. A basic decomposition result for isotropic distributions is this. Let $U \in R^n$ have the uniform distribution, say σ_n, on S_n. Thus σ_n is the unique $\mathcal{O}_n$-invariant probability measure on S_n. Consider a random variable $R \in (0, \infty)$ which is independent of U. It is clear that the distribution of $X = RU$ is isotropic. Conversely, if $\mathscr{L}(X) \in D_n$, then $U = X/\|X\|$ has the distribution σ_n on S_n and U is independent of $R = \|X\| \in (0, \infty)$. This well-known result is easily proved using the uniqueness of the $\mathcal{O}_n$-invariant probability measure σ_n on S_n.

For positive integers m and n, consider $U \in S_{n+m}$ which has σ_{n+m} as its distribution and let $U_{(m)} \in B_m$ denote the vector of the first m coordinates of U. The density function of $U_{(m)}$ (with respect to Lebesgue measure on R^m) at $u \in R^m$ is $\Psi(\|u\|^2 \mid m, n)$ where

$$\Psi(t \mid m, n) = c(m, n)(1 - t)^{(n/2)-1} I(t), \qquad t \in R^1. \tag{1}$$

In this formula,

$$c(m, n) = \Gamma\left(\frac{m+n}{2}\right) \Big/ \left[\left(\Gamma\left(\frac{1}{2}\right)\right)^m \Gamma\left(\frac{n}{2}\right)\right]$$

and $I(t) = 1$ if $0 < t < 1$ and $I(t) = 0$ otherwise. This result is easily proved by first observing the $\mathscr{L}(U) = \mathscr{L}(X/\|X\|)$ where X has coordinates $X_1, \cdots, X_{m+n}$ which are independent and identically distributed as $N(0, 1)$. Thus, we may assume that $U_{(m)}$ has coordinates $X_i/\|X\|$ for $i = 1, \cdots, m$. But, the joint distribution of $X_i^2/\|X\|^2$, $i = 1, \cdots, m$ is Dirichlet $D(½, \cdots, ½; n/2)$ (see Wilks (1962), page 177). Since the distribution of $U_{(m)}$ is invariant under sign changes of coordinates, the expression (1) for the density of $U_{(m)}$ follows by making the square root transformation on each coordinate in the expression for the Dirichlet density and multiplying by the appropriate Jacobian. Also, this argument shows that $\|U_{(m)}\|^2$ has a $\mathscr{B}(m/2, n/2)$ distribution—the beta distribution with parameters $m/2$ and $n/2$. For some related results, see Kingman (1963).

3. Marginals of isotropic distributions. Consider a random vector $X \in R^{m+n}$ and let $X_{(m)} \in R^m$ be the vector of the first m-coordinates of X. When $\mathscr{L}(X) \in D_{m+n}$, the distribution of $X_{(m)}$ is in D_m and this distribution will be called the *m-marginal* of $\mathscr{L}(X)$. In this section, we will give necessary and sufficient conditions that a distribution be the m-marginal of some distribution in D_{m+n}.

DEFINITION 2. The class of m-marginals of elements of D_{m+n} is denoted by $D(m, n)$.

REMARK. If $x_1, \cdots, x_{m+n}$ is any orthonormal basis for R^{m+n} and $\mathscr{L}(X) \in D_{m+n}$, let $\tilde{X}_{(m)} \in R^m$ have coordinates $x_i'X$, $i = 1, \cdots, m$. Then $\tilde{X}_{(m)}$ and $X_{(m)}$ have the same distribution since the distribution of X is isotropic. The standard orthonormal basis for R^{m+n} has been used for convenience.

THEOREM 1. *Let μ be a probability measure on R^m. The following are equivalent:*
(i) $\mu \in D(m, n)$;
(ii) *μ has a density with respect to Lebesgue measure, say f, given by $f(x) = h(\|x\|^2)$ where*

$$h(t) = \int_0^\infty \Psi\left(\frac{t}{r} \,\middle|\, m, n\right) \frac{1}{r^{m/2}} G(dr). \tag{2}$$

Here, $\Psi(\cdot \mid m, n)$ is given by (1) and G is any right continuous distribution function on R^1 with $G(0) = 0$.

PROOF. First, assume $\mu \in D(m, n)$ so there exists an $X \in R^{m+n}$ with $\mathscr{L}(X) \in D_{m+n}$ and $\mathscr{L}(X_{(m)}) = \mu$. Write $\mathscr{L}(X) = \mathscr{L}(RU)$ where $R \in (0, \infty)$ is independent of U which has a uniform distribution on S_{m+n}. Thus, $\mathscr{L}(X_{(m)}) = \mathscr{L}(RU_{(m)}) = \mu$. To show f is a density for μ, let $C \subseteq R^m$ be a Borel set and let H denote the distribution function of R. Now, compute as follows:

$$\begin{aligned}\mu(C) = P\{RU_{(m)} \in C\} &= \int_0^\infty \int_{R^m} I_C(ru)\Psi(\|u\|^2 \mid m, n)\, du\, H(dr)\\ &= \int_0^\infty \int_{R^m} I_C(u)\Psi\left(\frac{\|u\|^2}{r^2} \,\middle|\, m, n\right) \frac{1}{r^m}\, du\, H(dr)\\ &= \int_{R^m} I_C(u) \int_0^\infty \Psi\left(\frac{\|u\|^2}{r} \,\middle|\, m, n\right) \frac{1}{r^{m/2}} G(dr)\, du\\ &= \int_{R^m} I_C(u) f(u)\, du\end{aligned}$$

where G is the distribution function of the random variable $(R)^2$. Thus, f is the density of μ.

Conversely, suppose a measure μ on R^m has a density $f(x) = h(\|x\|^2)$ where h is given by (2). Let U, uniform on S_{m+n}, be independent of $R > 0$ and let $(R)^2$ have distribution G. Set $X = RU$ to see that $\mathscr{L}(X_{(m)}) = \mu$, so $\mu = \in D(m, n)$. This completes the proof.

Theorem 1 has some interesting implications. For notational convenience, let $\mathscr{G}$ be the set of distributions on R^1 with $G(0) = 0$. First, note that the representing distribution G in (2) is unique—that is, if

$$\int_0^\infty \Psi\left(\frac{t}{r} \,\middle|\, m, n\right) \frac{1}{r^{m/2}} G_1(dr) = \int_0^\infty \Psi\left(\frac{t}{r} \,\middle|\, m, n\right) \frac{1}{r^{m/2}} G_2(dr) \tag{3}$$

for all $t > 0$, then $G_1 = G_2$. To see this, the results in Williamson (1956) (see especially Theorem 7) show that for $\alpha > 0$ and $c > 0$, there exists a nondecreasing function γ_c such that

$$\int_0^{1/x} (1 - ux)^{\alpha-1}\gamma_c(du) = \exp[-cx]$$

for $x > 0$. Williamson's results are only stated for $\alpha \geq 1$ but can be extended to the case $\alpha > 0$. Taking $\alpha = n/2$ and integrating both sides of (3) with respect to $\gamma_c(dt)$ implies that

$$\int_0^\infty \exp[-c/r] \frac{1}{r^{m/2}} G_1(dr) = \int_0^\infty \exp[-c/r] \frac{1}{r^{m/2}} G_2(dr)$$

for all $c > 0$. The uniqueness of Laplace transforms implies that $G_1 = G_2$.

This uniqueness and Theorem 1 imply that the set of extreme points of the convex set $D(m, n)$ consists exactly of the probability measures μ_r, $r > 0$ where

$$\mu_r(B) = \int_B \Psi\left(\frac{\|u\|^2}{r} \,\middle|\, m, n\right) \frac{1}{r^{m/2}}\, du$$

for any Borel set $B \subseteq R^m$. An alternative way to phrase Theorem 1 is that $\mu \in D(m, n)$ if and only if

$$\mu = \int_0^\infty \mu_r G(dr)$$

for some $G \in \mathscr{G}$. This is a Choquet-type representation for the elements of $D(m, n)$.

By definition, it is clear that $D(m, n) \supseteq D(m, n+1)$. Define $D(m, \infty)$ by $D(m, \infty) = \cap_{m=1}^{\infty} D(m, n)$. Let $N(0, I_m)$ denote the m-dimensional normal distribution with mean zero and covariance I_m. The density of $N(0, I_m)$ will be denoted by

$$\varphi_m(u) = \frac{1}{(\sqrt{2\pi})^m} \exp\left(-\frac{1}{2}\|u\|^2\right), \qquad u \in R^m.$$

THEOREM 2. *The following are equivalent for a probability measure μ on R^m:*
(i) $\mu \in D(m, \infty)$;
(ii) *μ has a density f given by*

$$f(u) = \int_0^\infty \frac{1}{r^{m/2}} \varphi_m\left(\frac{u}{\sqrt{r}}\right) G(dr)$$

where $G \in \mathscr{G}$.

PROOF. If (ii) holds, then $\mu = \mathscr{L}(RZ_{(m)})$ where $\mathscr{L}(Z_{(m)}) = N(0, I_m)$, $\mathscr{L}[(R)^2] = G$ and R is independent of $Z_{(m)}$. Given n, let Z be independent of R with $\mathscr{L}(Z) = N(0, I_{m+n})$. Obviously $\mathscr{L}(RZ) \in D_{m+n}$ and $\mu = \mathscr{L}(RZ_{(m)}) = \mathscr{L}((RZ)_{(m)})$ so $\mu \in D(m, n)$ for all n. Hence (i) holds.

Conversely, suppose $\mu \in D(m, n)$ for all $n \geq 1$ so μ has a density $f(u) = h(\|u\|^2)$ where

$$h(t) = \int_0^\infty \Psi\left(\frac{t}{r} \,\middle|\, m, n\right) \frac{1}{r^{m/2}} \tilde{G}_n(dr).$$

For each t, it is easy to show

$$\lim_{n\to\infty} \frac{1}{n^{m/2}} \Psi\left(\frac{t}{n} \,\middle|\, m, n\right) = \frac{1}{(\sqrt{2\pi})^m} e^{-t/2}.$$

By Scheffé's theorem (see Billingsley (1968), page 224)

$$J_n(w) \equiv \int_{R^m} e^{iw'u} \frac{1}{n^{m/2}} \Psi\left(\frac{\|u\|^2}{n} \,\middle|\, m, n\right) du$$

converges uniformly to $\exp[-\frac{1}{2}\|w\|^2]$ for $w \in R^m$. Let $G_n(dr) = \tilde{G}_n(dr/n)$. Then for each $w \in R^m$,

$$\begin{aligned}
\xi(w) &\equiv \int_{R^m} e^{iw'u} h(\|u\|^2)\, du \\
&= \int_{R^m} \int_0^\infty e^{iw'u} \frac{1}{(nr)^{m/2}} \Psi\left(\frac{\|u\|^2}{nr} \,\middle|\, m, n\right) G_n(dr)\, du \\
&= A_n(G_n) + \int_0^\infty \exp\left[-\frac{1}{2} r \|w\|^2\right] G_n(dr)
\end{aligned}$$

where

$$A_n(G_n) = \int_0^\infty [J_n(\sqrt{r}w) - e^{-1/2r\|w\|^2}]G_n(dr).$$

Since $J_n(\cdot)$ converges to $e^{-1/2\|\cdot\|^2}$ uniformly, $\lim |A_n(G_n)| = 0$ for any sequence $\{G_n\}_{n=1}^\infty$. Let G_∞ be a weak limit point of the sequence $\{G_n\}$. Since $e^{-1/2r\|w\|^2}$ is a bounded continuous function of r, we conclude that

$$\xi(w) = \int_0^\infty e^{-1/2r\|w\|^2} G_\infty(dr)$$

for each $w \in R^m$. Since $\xi(0) = 1$, G_∞ is a proper distribution function and it is easy to show $G_\infty(0) = 0$. The uniqueness of characteristic functions implies that (ii) holds, so the proof is complete.

When $m = 1$, the result of Theorem 2 is well known and has been proved in several different, but equivalent, forms. For example, see Schoenberg (1938), Freedman (1963), and Kingman (1973). Other results concerning scale mixtures of the univariate normal distribution are given in Teichroew (1957), Kudo (1963), Kelker (1971), Andrews and Mallows (1974), and Efron and Olshen (1978).

4. A derivative condition. In this section, we present an alternative characterization of the functions h satisfying (2). This description is based on the similarity of (2) with Abel's integral equation, fractional derivatives and integrals (see Zygmund (1968), page 133), and the results given in Williamson (1956). Given a parameter $\alpha > 0$, a function f defined on $(0, \infty)$ is *α-times monotone* (see Williamson (1956), Definition 3 where the case of $\alpha \geq 1$ is treated) if

$$f(t) = \int_0^\infty (1 - ut)_+^{\alpha-1}\gamma(du)$$

where γ is nondecreasing with $\gamma(0) = 0$. Here, $(v)_+^{\alpha-1}$ for $v \in R$ is defined to be $v^{\alpha-1}$ if $v > 0$ and zero otherwise. In this notation, Theorem 1 shows that $\mu \in D(m, n)$ iff μ has a density $h(\|x\|^2)$ which lies in the subclass of $(n/2)$-times monotone functions for which

$$h(t) = \int_0^\infty (1 - ut)_+^{(n/2)-1}\gamma(du)$$

and

$$\gamma\left(d\left(\frac{1}{r}\right)\right) = c(m, n)\frac{1}{r^{m/2}}G(dr)$$

with $G \in \mathcal{G}$. Note that $\gamma(\infty) < +\infty$ iff $h(0) < +\infty$. When n is even, $n = 2, 4, \cdots$ we have the following

PROPOSITION 1. *The measure μ is in $D(m, n)$* iff *μ has a density $h(\|x\|^2)$ where*:
(i) *if $n = 2$, h is nonincreasing*;
(ii) *if $n/2 \equiv j \geq 2$, $(-1)^{j-2}h^{(j-2)}(t)$ is nonincreasing and convex.*

PROOF. This is an easy consequence of Theorem 4 in Williamson (1956) and the fact that $h(\|x\|^2)$ is a density on R^m.

When $n \equiv 2j + 1$ is odd, $j = 0, 1, 2, \cdots$, the situation is slightly different. For $j = 0$, $\mu \in D(m, 1)$ iff the density of μ, say $h(\|x\|^2)$ for $x \in R^m$, exists and has the form

$$h(t) = c(m, 1)\int_0^\infty \left(1 - \frac{t}{r}\right)_+^{-1/2} \frac{1}{r^{m/2}} G(dr) = \int_0^\infty (r - t)_+^{-1/2}\Psi(dr)$$

where

$$\Psi(r) = -c(m, 1) \int_r^\infty \frac{1}{(m-1)/u^2} G(du)$$

is nondecreasing with $\Psi(+\infty) = 0$. An easy application of Tonelli's theorem (see Dunford and Schwartz (1958), page 194) shows that

$$Q(s) \equiv \int_0^\infty (t-s)_+^{-1/2} h(t)\, dt = -\pi\Psi(s), \qquad s > 0$$

is nonincreasing with $Q(+\infty) = 0$. Conversely, suppose we are given h defined on $(0, \infty)$ to $[0, \infty)$ such that $h(\|x\|^2)$ is a density on R^m and

$$Q(s) \equiv \int_0^\infty (t-s)_+^{-1/2} h(t)\, dt, \qquad s > 0$$

is nonincreasing with $Q(+\infty) = 0$. Define h_0 by

$$h_0(t) = \int_0^\infty (r-t)_+^{-1/2} \Psi(dr)$$

where $\Psi(r) = -Q(r)/\pi$. Applying Tonelli's theorem again shows

$$Q(s) = \int_0^\infty (t-s)_+^{-1/2} h_0(t)\, dt, \qquad s > 0$$

and by the uniqueness of the above transform it follows that $h = h_0$ a.e. on $(0, \infty)$. From Theorem 1 we have that the measure μ defined by h on R^m is in $D(m, 1)$ iff $Q(s)$ is nonincreasing with $Q(+\infty) = 0$.

Now, consider the case when $n = 2j + 1$ is odd and $j \geq 1$.

PROPOSITION 2. *Let $h \geq 0$ be defined on $(0, \infty)$ such that $h(\|x\|^2)$ is a density on R^m yielding the probability measure μ. For $n = 2j + 1, j = 1, 2, \cdots, \mu \in D(m, n)$ iff the function*

$$Q(s) \equiv \int_0^\infty (t-s)_+^{-1/2} h(t)\, dt$$

satisfies $(-1)^{j-1} Q^{(j-1)}(s)$ is nonincreasing and convex with $Q(+\infty) = 0$.

PROOF. This follows by applying Williamson's Theorem 4 to the function Q. The details are left to the reader.

5. *n*-dimensional versions. In this section we give a partial answer to Q1 described earlier. Consider a real valued random variable Z with a symmetric distribution such that $P\{Z = 0\} = 0$—that is, $\mathscr{L}(Z) \in D_1$.

DEFINITION 3. Given an integer $n > 1$, the distribution of a random vector $X \in R^n$ is called an *n-dimensional version* of $\mathscr{L}(Z)$ if there exists a function c on R^n to $[0, \infty)$ such that

(i) $c(a) > 0$ if $a \neq 0$;

(ii) $\mathscr{L}(a'X) = \mathscr{L}(c(a)Z), a \in R^n$.

In what follows, we will often call X an n-dimensional version of Z when $\mathscr{L}(X)$ is an n-dimensional version of $\mathscr{L}(Z)$. If X is an n-dimensional version of Z such that $\mathscr{L}(\Gamma X) =$

$\mathscr{L}(X)$ for all $\Gamma \in \mathcal{O}_n$, then X will be called an *n-dimensional isotropic version* of Z. In Definition 3, the condition that $c(a) > 0$ for $a \neq 0$ is simply to guarantee that X be n-dimensional—that is, $P(X \in M) = 0$ for all proper subspaces M of R^n. With $m = 1$ in Theorem 1, we have a complete description of those distributions having n-dimensional isotropic versions.

DEFINITION 4. The set of all distributions in D_1 which have n-dimensional versions will be denoted by $\mathscr{F}_n$, $n \geq 2$.

It is obvious that $D(1, n-1) \subseteq \mathscr{F}_n$, $n = 2, \cdots$. Our first result identifies those distributions in $\mathscr{F}_n$ which have a finite variance.

THEOREM 3. *Suppose $\mathscr{L}(Z) \in \mathscr{F}_n$ and $\operatorname{Var}(Z) < +\infty$. Then $\mathscr{L}(Z) \in D(1, n-1)$ and every n-dimensional version of Z is given by AX_0 where X_0 is an n-dimensional isotropic version of Z and A is an $n \times n$ nonsingular matrix.*

PROOF. Let X be an n-dimensional version of Z with $c: R^n \to [0, \infty)$ satisfying Definition 2. Since $\operatorname{Var}(Z) \equiv \sigma^2 < \infty$, it follows that X has a covariance matrix, say $\Sigma = \operatorname{Cov}(X)$. The relation $\mathscr{L}(a'X) = \mathscr{L}(c(a)Z)$ implies $c(a) = (a'\Sigma a)^{1/2}/\sigma$, $a \in R^n$ which shows that Σ is positive definite as $c(a) > 0$ for $a \neq 0$. Set $X_0 = \sigma\Sigma^{-1/2}X$ where $\Sigma^{-1/2}$ denotes the inverse of the positive definite square root of Σ. For $a \in R^n$,

$$\mathscr{L}(a'X_0) = \mathscr{L}((\sigma\Sigma^{-1/2}a)'X) = \mathscr{L}(c(\sigma\Sigma^{-1/2}a)Z) = \mathscr{L}(\| a \| Z)$$

which implies that X_0 is an n-dimensional isotropic version of Z. Thus $\mathscr{L}(Z) \in D(1, n-1)$ and, of course, the distribution of X_0 is unique as $\mathscr{L}(X_0) \in D_n$ and $\mathscr{L}(a'X_0) = \mathscr{L}(\| a \| Z)$. Set $A = (1/\sigma)\Sigma^{1/2}$ to complete the proof.

Theorem 3 identifies all those elements in $\mathscr{F}_n$ with a finite variance and Theorem 1 gives a representation for the densities of such distributions. Further, every n-dimensional version of finite variance distributions in D_1 is equivalent (up to a linear transformation) to an isotropic n-dimensional version. However, the situation is more complicated in the infinite variance case. Let Z_α denote a symmetric stable law of order α with characteristic function

$$\eta_\alpha(t) = \exp[-|t|^\alpha], \qquad t \in R^1,$$

where $0 < \alpha < 2$. Given n and a probability measure δ on S_n such that $\delta(B) = \delta(-B)$ for Borel sets B, let

$$l_\delta(w) = \int_{S_n} |u'w|^\alpha \, \delta(du).$$

For each fixed α, the function

$$\xi(w) = \exp[-l_\delta(w)]$$

is a characteristic function of a symmetric n-dimensional stable law of order α (see Fristedt (1972)). Further, if $X \in R^n$ has characteristic function ξ, then

$$\mathscr{L}(a'X) = \mathscr{L}(c_\alpha(a)Z_\alpha)$$

where

$$c_\alpha(a) = \left[\int_{S_n} |u'a|^\alpha \, \delta(du)\right]^{1/\alpha}.$$

Thus, X is an n-dimensional version of Z_α for each δ and X is isotropic iff $\delta = \sigma_n$. Further, different δ's yield nonequivalent (up to linear transformations) n-dimensional versions of Z_α. Of course, $\mathscr{L}(Z_\alpha) \in D(1, n-1)$ for all n—see Andrews and Mallows (1974) for a representation of $\mathscr{L}(Z_\alpha)$ as a scale mixture of normals.

6. Remarks. It is clear that $\mathscr{F}_n \supseteq \mathscr{F}_{n+1}$ so $\mathscr{F}_\infty = \lim_{n\to\infty} \mathscr{F}_n$ is well defined. The obvious question is whether or not $\mathscr{F}_n = D(1, n-1)$ and, in particular, whether or not $\mathscr{F}_\infty = D(1, \infty)$. This latter equality would establish that the only symmetric distributions which have n-dimensional versions for all n are scale mixtures of the normal distribution.

In the case when $\mathrm{Var}(Z) < +\infty$, Theorem 3 shows that the only scale functions $c: R^n \to [0, \infty)$ are quadratic. More precisely, if X is any n-dimensional version of Z with $\mathrm{Var}(Z) < +\infty$, then the function c is $c(a) = a'\Sigma a$, $a \in R^n$, for some positive definite Σ. However, in the infinite variance case, other scale functions can arise such as the c_α discussed in the previous section. One possible approach to the questions mentioned in the first paragraph is to try to characterize the scale functions which can arise. This author has made no significant progress in answering any of the questions raised above.

We close this section with a few comments about the distributions in $D(1, \infty)$. If Z_1 and Z_2 are independent with $\mathscr{L}(Z_i) \in D(1, \infty)$, $i = 1, 2$, it is not difficult to show that $\mathscr{L}(Z_1 + Z_2) \in D(1, \infty)$, $\mathscr{L}(Z_1Z_2) \in D(1, \infty)$, and $\mathscr{L}(Z_1/Z_2) \in D(1, \infty)$. However, $\mathscr{L}(Z) \in D(1, \infty)$ does not imply that $\mathscr{L}(1/Z) \in D(1, \infty)$.

When f is a symmetric density on R which yields a distribution in $D(1, \infty)$, Theorem 2 provides the representation

$$f(u) = \int_0^\infty \frac{1}{\sqrt{2\pi r}} \exp\left[-\frac{u^2}{2r}\right] G(dr).$$

If

$$f(0) = \frac{1}{\sqrt{2\pi}} \int_0^\infty \frac{1}{\sqrt{r}}\, G(dr) < +\infty,$$

it is not difficult to show that

$$\hat{f}(t) \equiv \int_{-\infty}^\infty e^{itu} f(u)\, du$$

is nonnegative and

$$\int_{-\infty}^\infty \hat{f}(t)\, dt = 2\pi f(0).$$

From this it follows that the distribution determined by the density $\hat{f}/2\pi f(0)$ is in $D(1, \infty)$. In particular, $k(\alpha)\exp[-|u|^\alpha]$ is a density (for a proper choice of $k(\alpha)$) which yields a distribution in $D(1, \infty)$ for $0 < \alpha \le 2$.

7. *Postscript to Section 2.* The following, suggested by Persi Diaconis and David Freeman, provides an alternative proof of Theorem 1 and the uniqueness of G in the representation of the elements of $D(m, n)$. Let $\sigma_{n,r}$ denote the uniform distribution on $S_{n,r} = \{x \mid x \in R^n, \|x\| = r\}$ where $r > 0$. The discussion of Section 2 shows that a probability measure μ is in $\mathfrak{D}_n$ iff

$$\mu = \int_0^\infty \sigma_{n,r}\, G(dr) \tag{4}$$

for some $G \in \mathscr{G}$. The representing G is unique since if $\mathscr{L}(X) = \mu$ and μ is given by (3), then $\mathscr{L}(\|X\|) = G$. This shows that the set of extreme points of D_n is the set $\{\sigma_{n,r} \mid r > 0\}$. Given positive integers m and n, define the affine transformation $T_{m,n}$ on D_{m+n} to $D(m, n)$ by

$$(T_{m,n}\mu)\ (B) = \mu\ (B \times R^n)$$

where B is a Borel subset of R^m. Thus, if $\mathscr{L}(X) = \mu \in D_{m+n}$, then $T_{m,n}\mu$ is the distribution of the first m coordinates of X. By definition, $D(m, n)$ is the image of D_{m+n} under the mapping $T_{m,n}$. Here is Theorem 1 in the present notation. The assertion is that $\nu \in D(m, n)$ iff

$$\nu = \int_0^\infty T_{m,n}\ (\sigma_{m+n,r}) G(dr) \tag{5}$$

for some $G \in \mathscr{G}$. To show this, if $\nu \in D(m, n)$, then $\nu = T_{m,n}\mu$ for some $\mu \in D_{n,n}$. Using the representation (4) for μ and the fact that $T_{m,n}$ is affine, we have

$$\nu = T_{m,n}\mu = \int_0^\infty T_{m,n}\ (\sigma_{m+n,r}) G(dr). \tag{6}$$

Conversely, if ν is given by (5), then $\nu = T_{n,n}\mu$ where μ is given by (4) with the same G which gives ν in (5).

The uniqueness of G in (5) will follow once we show that $T_{m,n}$ is a one-to-one function. First consider $T_{l,k}$ on D_{l+k} to $D(l,k)$. If $\mathscr{L}(X) = \mu \in D_{l+k}$, then $T_{l,k}\mu$ is the probability distribution of the first coordinate of X-say X_l. Thus, from $T_{l,k}\mu$, we can calculate the characteristic function of X_l-say η. However, the $\mathcal{O}_n$-invariance of μ implies that the characteristic function of X is given by

$$\mathscr{E}(e^{it'X}) = \eta(\|t\|)$$

for $t \in R^{l+k}$. Therefore, the measure μ is determined by the measure $T_{l,k}\mu$ so $T_{l,k}$ is one-to-one. That $T_{m,n}$ is one-to-one is a consequence of the equation

$$T_{l,m+n-1} = T_{l,m} \circ T_{m,n}$$

and the fact that $T_{l,m+n-l}$ is one-to-one. The uniqueness of G in (5) now follows from the uniqueness of G in (4) and the observation that $T_{m,n}$ is one-to-one. This implies that $\{T_{m,n}(\sigma_{m+n,r}) \mid r > 0\}$, is exactly the set of extreme points of the convex set $D(m, n)$.

REFERENCES

ANDREWS, D. F. and MALLOWS, C. L. (1974). Scale mixtures of normal distributions. *J. Roy. Statist. Soc. Ser. B* **36** 99–102.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

DUNFORD, N. and SCHWARTZ, J. T. (1958). *Linear Operators, Part I.* Interscience, New York.

EATON, M. L. (1977). N-dimensional versions of some symmetric univariate distributions. Univ. Minn. Technical Report No. 288.

EFRON, B. and OLSHEN, R. (1978). How broad is the class of normal scale mixtures? *Ann. Statist.* **6** 1159–1164.

FREEDMAN, D. (1963). Invariants under mixing which generalize deFinetti's theorem: Continuous time parameter. *Ann. Math. Statist.* **34** 1194–1226.

FRISTEDT, B. (1972). Expansions for the density of the absolute value of a strictly stable vector. *Ann. Math. Statist.* **43** 669–672.

KELKER, D. (1971). Infinite divisibility and variance mixtures of the normal distribution. *Ann. Math. Statist.* **42** 802–808.

KINGMAN, J. F. C. (1963). Random walks with spherical symmetry. *Acta. Math.* **109** 11–53.

KINGMAN, J. F. C. (1972). On random sequences with spherical symmetry. *Biometrika* **59** 492–494.

KUDO, A. (1963). A multivariate analogue of the one-sided t-test. *Biometrika* **50** 403–418.

SAVAGE, L. J. (1969). Notes on round distributions. Leonard Savage papers, Yale Univ. Library, New Haven, Connecticut.

SCHOENBERG, I. J. (1938). Metric spaces and completely monotone functions. *Ann. Math.* **39** 811–841.
TEICHROEW, D. (1957). The mixture of normal distributions with different variances. *Ann. Math. Statist.* **28** 510–512.
WILKS, S. (1962). *Mathematical Statistics.* Wiley, New York.
WILLIAMSON, R. E. (1956). Multiply monotone functions and their Laplace transforms. *Duke Math. J.* **23** 189–207.
ZYGMUND, A. (1968). *Trigonometric Series, Vol. II.* Cambridge Univ.

DEPT. OF THEORETICAL STATISTICS
UNIVERSITY OF MINNESOTA
270a VINCENT HALL
206 CHURCH STREET S.E.
MINNEAPOLIS, MINNESOTA 55455