SOME CHARACTERIZATION PROBLEMS IN PROBABILITY THEORY

By

W. LYLE COOK

Bachelor of Science

Southeast Missouri State College

Cape Girardeau, Missouri

1965

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY
July, 1972

SOME CHARACTERIZATION PROBLEMS IN

PROBABILITY THEORY

Thesis Approved:

Kotlarski

Thesis Adviser

A. H. Som

John Jewett

Joe L. Howard

D. Durham

Dean of the Graduate College

ACKNOWLEDGMENTS

I would like to express my appreciation to Dr. Ignacy Kotlarski for serving as my research adviser and for his patient personal guidance. My sincerity here cannot be overemphasized. I would also like to thank the other members of my committee, Dr. John Jewett, Dr. Joe Howard, and Dr. A. H. Soni, for their assistance.

I would like to thank Mrs. Margaret Estes for her superb typing.

Finally, I would like to thank my wife, LaVerne, whose sacrifice during the past few years is appreciated. She may rest assured that years of pleasant homemaking lie ahead of her.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

The problem of characterizing probability distribution originates in early papers by Bernstein [2] and Cramer [3]. The modern equivalents of the results of these men are stated in the following.

Theorem 1.1 (Bernstein): Let X_1 and X_2 be independent random variables. Then the independence of $Z_1 = X_1 + X_2$ and $Z_2 = X_1 - X_2$ is a necessary and sufficient condition for X_1 and X_2 to be distributed normally.

The result of Cramer is the converse to the well known fact that if X_1 and X_2 are independent and normally distributed then so is the sum $X_1 + X_2$.

Theorem 1.2 (Cramer): Let X_1 and X_2 be independent random variables and suppose $Z = X_1 + X_2$ is distributed normally. Then each of X_1 and X_2 is distributed normally.

We note that Bernstein's characterization is based on functions of random variables being independent, while Cramer's relies on a property of a statistic, in particular that of being normally distributed.

It is the purpose of this paper to extend and generalize two characterizations, one given originally by M. V. Tamhankar [10], the second by A. Kagan and O. Shalaevski [5]. The work with Tamhankar's characterization is similar to the theorem of Bernstein in that it is based on independent functions. The generalizations of the work by Kagan and Shalaevski are based on statistic properties as was Cramer's theorem. These later results are presented as they occurred in our investigations, and thus tend to telescope.

We state the theorem of Tamhankar in nearly the form in which it originally appeared. Let $X_1, \ldots, X_n$ have a joint probability density function $f(x_1, \ldots, x_n)$ and transform as follows:

$$
\begin{aligned}
X_1 &= R \cos \Theta_1 \\
X_2 &= R \sin \Theta_1 \cos \Theta_2 \\
&\vdots \\
X_{n-1} &= R \sin \Theta_1 \cdot \ldots \cdot \sin \Theta_{n-2} \cos \Theta_{n-1} \\
X_n &= R \sin \Theta_1 \cdot \ldots \cdot \sin \Theta_{n-2} \sin \Theta_{n-1}
\end{aligned}
$$

where capital letters represent random variables and small letters their values and

$$
\begin{aligned}
&-\infty < x_i \leq \infty && i = 1, 2, \ldots, n \\
&0 \leq r < \infty && \\
&0 \leq \theta_i \leq \pi && i = 1, 2, \ldots, n-2 \\
&0 \leq \theta_{n-1} < 2\pi &&
\end{aligned}
$$

Let g be the probability density function of the transformed vector $(R, \Theta_1, \ldots, \Theta_{n-1})$ and assume $g(r, \theta_1, \ldots, \theta_{n-1})/|J|$ to be well defined and continuous so we may write

$$f(x_1, \ldots, x_n)|J| = g(r, \theta_1, \ldots, \theta_{n-1})$$

where

$$J = \frac{\partial(x_1, \ldots, x_n)}{\partial(r, \theta_1, \ldots, \theta_{n-1})} = r^{n-1} \sin^{n-2} \theta_1 \cdot \ldots \cdot \sin \theta_{n-2} .$$

Assume also that f is continuous in each X_j and note that $r^2 = \sum_{j=1}^{n} x_j^2$.

Theorem 1.3 (Tamhankar): Under the above conditions $X_1, \ldots, X_n$ are mutually independent and R is independent of $(\Theta_1, \ldots, \Theta_{n-1})$ if and only if all X_j are distributed normally with zero mean and common variance.

The theorem was extended by Kotlarski [6, 7] to allow more general transformations, and by Flusser [4] to allow more freedom in the independence grouping of the original and transformed vectors; i.e., $(R, \ldots, \Theta_j)$ and $(\Theta_{j+1}, \ldots, \Theta_{n-1})$. In Chapter II we combine and generalize the work of Kotlarski and Flusser to give what appears to be the ultimate extension of Tamhankar's result.

It is well known that if the statistic $Z = \sum_{j=1}^{n} (X_j + a_j)^2$, $a_j \in R$, is drawn from a population which is distributed normally with zero mean, then the distribution of Z depends on the a_j only through $\sum_{j=1}^{n} a_j^2$. The theorem of Kagan and Shalaevski is the converse.

Theorem 1.4 (Kagan and Shalaevski): Let $X_1, \ldots, X_n$ be independent and identically distributed (i.i.d.) and suppose the distribution of $Z = \sum_{j=1}^{n} (X_j + a_j)^2$, $a_j \in R$, depends on the a_j only through $\sum_{j=1}^{n} a_j^2$. Then each X_j is distributed $N(0,\sigma)$.

In Chapter III we present three generalizations of this theorem. The first relaxes the i.i.d. requirement to the independence of vectors $(X_1, \ldots, X_m)$ and $(X_{m+1}, \ldots, X_n)$ with no restriction that the variates have an identical distribution. The second allows the characterization of correlated random variables and the third invokes convolution and Fourier transforms to characterize other probability distributions, including the gamma.

CHAPTER II

GENERALIZATION OF TAMHANKAR'S THEOREM

Let $\vec{X} = (\vec{W}, \vec{Z})$, where $\vec{W} = (X_1, \ldots, X_s)$ and $\vec{Z} = (X_{s+1}, \ldots, X_n)$, be an absolutely continuous random vector. Let $\vec{W}$ and $\vec{Z}$ be independent with continuous positive density functions as follows:

(C-1) $\vec{W}$: $f(x_1, \ldots, x_s)$, where $(x_1, \ldots, x_s) \in \mathcal{G}_1 \subset R^s$ with $(0, \ldots, 0) \in \mathcal{G}_1$. $\quad 1 < s < n$

$\vec{Z}$: $g(x_{s+1}, \ldots, x_n)$, where $(x_{s+1}, \ldots, x_n) \in \mathcal{G}_2 \subset R^{n-s}$ with $(0, \ldots, 0) \in \mathcal{G}_2$.

Note that the Lebesque measure of $\mathcal{G}_1$ and $\mathcal{G}_2$ should be positive.

(C-2) Let p_k, $k = 1, \ldots, n$, be real numbers such that there exist limits

$$\lim_{x_1, \ldots, x_s \to 0} \frac{f(x_1, \ldots, x_s)}{\prod_{k=1}^{s} |x_k|^{p_k - 1}} = A_1 > 0 ,$$

$$\lim_{x_{s+1}, \ldots, x_n \to 0} \frac{g(x_{s+1}, \ldots, x_n)}{\prod_{k=s+1}^{n} |x_k|^{p_k - 1}} = A_2 > 0 .$$

Consider the transformation

$$
\begin{aligned}
x_1 &= \alpha_1(y_1)\alpha_2(y_2)\alpha_3(y_3) \;\ldots\; \alpha_r(y_r)\gamma(y_{r+1}, \ldots, y_n) \\
x_2 &= \alpha_1(y_1)\beta_2(y_2)\alpha_3(y_3) \;\ldots\; \alpha_r(y_r)\gamma(y_{r+1}, \ldots, y_n) \\
x_3 &= \alpha_1(y_1) \qquad \beta_3(y_3) \;\ldots\; \alpha_r(y_r)\gamma(y_{r+1}, \ldots, y_n) \\
&\;\;\vdots \\
x_r &= \alpha_1(y_1) \qquad\qquad \ldots\; \beta_r(y_r)\gamma(y_{r+1}, \ldots, y_n) \\
x_{r+1} &= \alpha_1(y_1) \qquad\qquad \ldots\; \gamma_{r+1}(y_{r+1}, \ldots, y_n) \\
&\;\;\vdots \\
x_n &= \alpha_1(y_1) \qquad\qquad \ldots\; \gamma_n(y_{r+1}, \ldots, y_n)
\end{aligned}
\tag{2.1.1}
$$

for $y_1 \in [0,\infty]$ and $(y_2, \ldots, y_n) \in \mathfrak{D}_0 \subset R^{n-1}$, with $1 \le r \le n$. Assume that the functions α_k, β_k, γ, and γ_k are taken in such a way that there is possible a change of variables in n-dimensional integrals, in particular that the Jacobian of (2.1.1) exists and does not vanish for $y_1 \in (0,\infty)$, $(y_2, \ldots, y_n) \in \mathfrak{D}_0$. We also require at least one point $(y^*_{r+1}, \ldots, y^*_n)$ such that $\gamma(y^*_{r+1}, \ldots, y^*_n) = 0$, and that α_1 be strictly monotonic on $[0,\infty)$ with $\alpha_1(0) = 0$.

(C-3) Assume there exist continuous real valued functions Ψ, Ψ_1, Ψ_2 defined on $[0,\infty)$, $\mathcal{G}_1$, $\mathcal{G}_2$, respectively, such that

$$\Psi(y_1) = \Psi_1(x_1, \ldots, x_s) + \Psi_2(x_{s+1}, \ldots, x_n)$$

$$y_1 \in [0, \infty), \quad (x_1, \ldots, x_s \in G_1$$

$$(x_{s+1}, \ldots, x_n) \in G_2 \qquad (2.1.2)$$

where each function maps onto $[0, \infty)$, has value 0 only at the origin, and Ψ is strictly monotonic.

(C-4) Finally, assume the functions

$$f(x_1, \ldots, x_s) = \begin{cases} A_1 \prod_{k=1}^{s} |x_k|^{p_k - 1} e^{-a\Psi_1(x_1, \ldots, x_s)} & (x_1, \ldots, x_s) \in G_1 \\ 0 & (x_1, \ldots, x_s) \in R^s \setminus G_1 \end{cases}$$

$$g(x_{s+1}, \ldots, x_n) = \begin{cases} A_2 \prod_{k=s+1}^{n} |x_k|^{p_k - 1} e^{-a\Psi_2(x_{s+1}, \ldots, x_n)} & (x_{s+1}, \ldots, x_n) \in G_2 \\ 0 & (x_{s+1}, \ldots, x_n) \in R^{n-s} \setminus G_2 \end{cases}$$

$$(2.1.3)$$

are probability density functions.

<u>Theorem 2.1</u>: $\vec{W}$ and $\vec{Z}$ are distributed as in (2.1.3) if and only if there exists an integer q, $1 \le q \le r$, such that $(Y_1, \ldots, Y_q)$ and $(Y_{q+1}, \ldots, Y_n)$ are independent.

We state the following lemmas, which may be verified through manipulation of the determinant, before proceeding with the proof of the theorem.

<u>Lemma 2.1</u>: The Jacobian of the transformation (2.1.1) is, for any $k \le r$, the product of two functions, one involving only $(y_1, \ldots, y_k)$ the other only $(y_{k+1}, \ldots, y_n)$. Thus

$$|J| = H_1(y_1, \ldots, y_k)\, H_2(y_{k+1}, \ldots, y_n)$$

$$1 \le k \le r \, . \tag{2.1.4}$$

<u>Lemma 2.2</u>: The product $\prod_{k=1}^{n} |x_k|^{p_k - 1}$ remains after applying transformation (2.1.1) a product of two functions, one involving only $(y_1, \ldots, y_q)$ the other only $(y_{q+1}, \ldots, y_n)$. Hence, we may write

$$\prod_{k=1}^{n} |x_k|^{p_k-1} = G(y_1, \ldots, y_q)\, H(y_{q+1}, \ldots, y_n) \, ,$$

$$(y_1, \ldots, y_q) \in \mathfrak{D}_1 \subset R^q$$

$$(y_{q+1}, \ldots, y_n) \in \mathfrak{D}_2 \subset R^{n-q} \, . \tag{2.1.5}$$

<u>Proof of Theorem</u>: Suppose q exists. The random vectors $(X_1, \ldots, X_n)$ and $(Y_1, \ldots, Y_n)$ have density functions f^* defined on $\mathfrak{C} = \mathfrak{C}_1 \times \mathfrak{C}_2$, and h^* defined on $\mathfrak{D} = [0,\infty) \times \mathfrak{D}_0$, respectively, which are connected by the formula

$$f^*(x_1, \ldots, x_n)|J| = h^*(y_1, \ldots, y_n) . \tag{2.1.6}$$

In view of the assumed independence of the variables this becomes

$$f(x_1, \ldots, x_s)\, g(x_{s+1}, \ldots, x_n)|J| = h(y_1, \ldots, y_q)\, k(y_{q+1}, \ldots, y_n), \tag{2.1.7}$$

$$\begin{cases} (x_1, \ldots, x_s) \in \mathfrak{C}_1 \\ (x_{s+1}, \ldots, x_n) \in \mathfrak{C}_2 \\ (y_1, \ldots, y_q) \in \mathfrak{D}_1 \subset R^q \\ (y_{q+1}, \ldots, y_n) \in \mathfrak{D}_2 \subset R^{n-q} \end{cases} \tag{2.1.8}$$

where $\mathfrak{D}_1 \times \mathfrak{D}_2 = \mathfrak{D}$, and the x's and y's are connected by (2.1.1).

Dividing both sides of (2.1.7) by $|J| = H_1 \cdot H_2$ and setting

$$\begin{cases} h_0(y_1, \ldots, y_q) = \dfrac{h(y_1, \ldots, y_q)}{H_1(y_1, \ldots, y_q)}, & (y_1, \ldots, y_n) \in \mathfrak{D}_1 \\ k_0(y_{q+1}, \ldots, y_n) = \dfrac{k(y_{q+1}, \ldots, y_n)}{H_2(y_{q+1}, \ldots, y_n)}, & (y_{q+1}, \ldots, y_n) \in \mathfrak{D}_2 \end{cases} \tag{2.1.9}$$

we have

$$f(x_1, \ldots, x_s)\, g(x_{s+1}, \ldots, x_n) = h_0(y_1, \ldots, y_q)\, k_0(y_{q+1}, \ldots, y_n) \tag{2.1.10}$$

with arguments as in (2.1.8).

Dividing side by side (2.1.5) into (2.1.10) yields

$$f_1(x_1, \ldots, x_s)\, g_1(x_{s+1}, \ldots, x_n) = h_1(y_1, \ldots, y_q)\, k_1(y_{q+1}, \ldots, y_n) \tag{2.1.11}$$

with arguments as in (2.1.8) and

$$\begin{cases} f_1(x_1, \ldots, x_s) = \dfrac{f(x_1, \ldots, x_s)}{\prod_{j=1}^{s} |x_j|^{p_j - 1}}, & (x_1, \ldots, x_s) \in \mathfrak{a}_1 \\[2ex] g_1(x_{s+1}, \ldots, x_n) = \dfrac{g(x_{s+1}, \ldots, x_n)}{\prod_{j=s+1}^{n} |x_j|^{p_j - 1}}, & (x_{s+1}, \ldots, x_n) \in \mathfrak{a}_2 \\[2ex] h_1(y_1, \ldots, y_q) = \dfrac{h_0(y_1, \ldots, y_q)}{G(y_1, \ldots, y_q)}, & (y_1, \ldots, y_q) \in \mathfrak{D}_1 \\[2ex] k_1(y_{q+1}, \ldots, y_n) = \dfrac{k_0(_{q+1}, \ldots, y_n)}{H(y_{q+1}, \ldots, y_n)}, & (y_{q+1}, \ldots, y_n) \in \mathfrak{D}_2 . \end{cases} \tag{2.1.12}$$

Letting $y_1 \to 0$ and observing from (2.1.1) that this forces each x_j to zero, $j = 1, \ldots, n$, causing (2.1.11) to become

$$f_1(0, \ldots, 0)\, g_1(0, \ldots, 0) = h_1(0, y_2, \ldots, y_q)\, k_1(y_{q+1}, \ldots, y_n)$$

$$(0, y_1, \ldots, y_q) \in \mathfrak{D}_1 , \quad (y_{q+1}, \ldots, y_n) \in \mathfrak{D}_2 . \tag{2.1.13}$$

We remark here that the left side of (2.1.13) is equivalent to the product of A_1 and A_2 by definition in (C-2). Now dividing (2.1.11) by (2.1.13) we obtain

$$f_2(x_1, \ldots, x_s)\, g_2(x_{s+1}, \ldots, x_n) = h_2(y_1, \ldots, y_q) \tag{2.1.14}$$

with arguments as in (2.1.8) and where

$$\begin{cases} f_2(x_1, \ldots, x_s) = \dfrac{f_1(x_1, \ldots, x_s)}{A_1} , & (x_1, \ldots, x_s) \in \mathfrak{A}_1 \\ g_2(x_{s+1}, \ldots, x_n) = \dfrac{g_1(x_{s+1}, \ldots, x_n)}{A_2} , & (x_{s+1}, \ldots, x_n) \in \mathfrak{A}_2 \\ h_2(y_1, \ldots, y_q) = \dfrac{h_1(y_1, \ldots, y_q)}{h_1(0, y_2, \ldots, y_q)} , & (y_1, \ldots, y_q) \in \mathfrak{D}_1 \; . \end{cases} \tag{2.1.15}$$

Obsere here that $f_2(0, \ldots, 0) = g_2(0, \ldots, 0) = 1$.

Choosing a vector $(y^*_{q+1}, \ldots, y^*_{r+1}, \ldots, y^*_n) \in \mathfrak{D}_2$ such that $\gamma(y^*_{r+1}, \ldots, y^*_n) = 0$ we note from (2.1.1) that at this point $x_1 = x_2 = \ldots = x_r = 0$. Since Equation (2.1.14) is independent of the choice of $(y_{q+1}, \ldots, y_n)$ it is equivalent to

$$\pi_2(x_{r+1}, \ldots, x_n) = h_2(y_1, \ldots, y_q), \quad (x_{r+1}, \ldots, x_n) \in R^{n-r}$$
$$(y_1, \ldots, y_q) \in \mathfrak{D}_1 \tag{2.1.16}$$

where π_2 is the product of f_2 and g_2 restricted to points at which the first r coordinates are zero. But observing again from (2.1.1) that the vector $(x_{r+1}, \ldots, x_n)$ is independent of $(y_2, \ldots, y_r)$ we are justified in writing

$$h_3(y_1) = h_2(y_1, y_2, \ldots, y_q), \; y_1 \in [0, \infty) \; , \; (y_1, \ldots, y_q) \in \mathfrak{D}_1 \; , \tag{2.1.17}$$

from which (2.1.14) becomes

$$f_2(x_1, \ldots, x_s) \, g_2(x_{s+1}, \ldots, x_n) = h_3(y_1), \; (x_1, \ldots, x_s) \in \mathfrak{A}_1 \; ,$$
$$(x_{s+1}, \ldots, x_n) \in \mathfrak{A}_2 \; , \quad y_1 \in [0, \infty) \; . \tag{2.1.18}$$

Since Ψ is one to one and onto, Ψ^{-1} exists and we may write (2.1.2) as

$$y_1 = \Psi^{-1}[\Psi_1(x_1, \ldots, x_s) + \Psi_2(x_{s+1}, \ldots, x_n)], \quad y_1 \in [0,\infty)$$

$$(x_1, \ldots, x_s) \in G_1, \quad (x_{s+1}, \ldots, x_n) \in G_2. \tag{2.1.19}$$

Then (2.1.18) becomes

$$f_2(x_1, \ldots, x_s)\; g_2(x_{s+1}, \ldots, x_n) = h_3[\Psi^{-1}(\Psi_1[x_1, \ldots, x_s] + \Psi_2[x_{s+1}, \ldots, x_n])],$$

$$(x_1, \ldots, x_s) \in G_1, \quad (x_{s+1}, \ldots, x_n) \in G_2. \tag{2.1.20}$$

If we set

$$h_3[\Psi^{-1}(u)] = h_4(u), \quad u \in [0,\infty) \tag{2.1.21}$$

the Equation (2.1.20) becomes

$$f_2(x_1, \ldots, x_s)\; g_2(x_{s+1}, \ldots, x_n) = h_4[\Psi_1(x_1, \ldots, x_s) + \Psi_2(x_{s+1}, \ldots, x_n)],$$

$$(x_1, \ldots, x_s) \in G_1, \quad (x_{s+1}, \ldots, x_n) \in G_2. \tag{2.1.22}$$

By evaluating (2.1.22) first when $x_1 = x_2 = \ldots = x_s = 0$ and then when $x_{s+1} = x_{s+2} = \ldots = x_n = 0$ we have

$$\begin{cases} f_2(x_1, \ldots, x_s) = h_4[\Psi_1(x_1, \ldots, x_s)], & (x_1, \ldots, x_s) \in G_1 \\ g_2(x_{s+1}, \ldots, x_n) = h_4[\Psi_2(x_{s+1}, \ldots, x_n)], & (x_{s+1}, \ldots, x_n) \in G_2 . \end{cases} \tag{2.1.23}$$

Putting (2.1.23) into (2.1.22)

$$h_4[\Psi_1(x_1, \ldots, x_s)]\, h_4[\Psi_2(x_{s+1}, \ldots, x_n)] =$$

$$= h_4[\Psi_1(x_1, \ldots, x_s) + \Psi_2(x_{s+1}, \ldots, x_n)],$$

$$(x_1, \ldots, x_s) \in G_1, \quad (x_{s+1}, \ldots, x_n) \in G_2 . \qquad (2.1.24)$$

Now set

$$\begin{cases} u_1 = \Psi_1(x_1, \ldots, x_s), & (x_1, \ldots, x_s) \in G_1 \\ u_2 = \Psi_2(x_{s+1}, \ldots, x_n), & (x_{s+1}, \ldots, x_n) \in G_2 \end{cases} \qquad (2.1.25)$$

from which (2.1.24) becomes

$$h_4(u_1)\, h_4(u_2) = h_4(u_1 + u_2), \quad u_k \in [0,\infty), \; k = 1,2 . \qquad (2.1.26)$$

The Equation (2.1.16) is the Cauchy equation and in view of the continuity of f, g, Ψ, Ψ_1, Ψ_2, the function h_4 is continuous. The solution of (2.1.26) [see [1] p. 38] is given by

$$h_4(u) = e^{-au}, \qquad u \in [0,\infty), \text{ and } a \text{ is a real constant.}$$

We may now retrace our steps. From (2.1.25)

$$\begin{cases} h_4[\Psi_1(x_1, \ldots, x_s)] = e^{-a\Psi_1(x_1, \ldots, x_s)}, & (x_1, \ldots, x_s) \in G_1 \\ h_4[\Psi_2(x_{s+1}, \ldots, x_n)] = e^{-a\Psi_2(x_{s+1}, \ldots, x_n)}, & (x_{s+1}, \ldots, x_n) \in G_2 . \end{cases}$$

From (2.1.23)

$$\begin{cases} f_2(x_1, \ldots, x_s) = e^{-a\Psi_1(x_1, \ldots, x_s)}, & (x_1, \ldots, x_s) \in G_1 \\ g_2(x_{s+1}, \ldots, x_n) = e^{-a\Psi_2(x_{s+1}, \ldots, x_n)}, & (x_{s+1}, \ldots, x_n) \in G_2 . \end{cases}$$

From (2.1.15)

$$\begin{cases} f_1(x_1, \ldots, x_s) = A_1 e^{-a\Psi_1(x_1, \ldots, x_s)}, & (x_1, \ldots, x_s) \in \mathfrak{a}_1 \\ g_1(x_{s+1}, \ldots, x_n) = A_2 e^{-a\Psi_2(x_{s+1}, \ldots, x_n)}, & (x_{s+1}, \ldots, x_n) \in \mathfrak{a}_2. \end{cases}$$

From (2.1.12)

$$\begin{cases} f(x_1, \ldots, x_s) = A_1 \prod_{k=1}^{s} |x_k|^{p_k - 1} e^{-a\Psi_1(x_1, \ldots, x_s)}, \\ \qquad (x_1, \ldots, x_s) \in \mathfrak{a}_1 \\ g(x_{s+1}, \ldots, x_n) = A_2 \prod_{k=s+1}^{n} |x_k|^{p_k - 1} e^{-a\Psi_2(x_{s+1}, \ldots, x_n)}, \\ \qquad (x_{s+1}, \ldots, x_n) \in \mathfrak{a}_2 . \end{cases}$$

This completes the proof of the sufficiency.

To see that the condition is necessary, note that when (2.1.3) holds the product of f and g may be substituted into (2.1.5). Recall that the Jacobian may be written as a product which when transformed by (2.1.1) and also used in (2.1.5) gives the desired separation of g^*.

It is possible to see that the theorems of Tamhankar [10] and Flusser [4] are corollaries of Theorem 2.1 as follows. Take $r = n$ and the functions α_k and β_k as given below:

$$\alpha_1(y_1) = y_1$$

$$\alpha_k(y_k) = \cos y_k, \qquad 2 \le k \le n$$

$$\beta_k(y_k) = \sin y_k, \qquad 2 \le k \le n$$

with $0 \le y_k \le \pi$ for $2 \le k \le n-1$ and $0 \le y_n < 2\pi$. Set

$$\Psi(y_1) = y_1^{\,2}$$

$$\Psi_1(x_1, \ldots, x_s) = x_1^{\,2} + \ldots + x_s^{\,2}$$

$$\Psi_2(x_{s+1}, \ldots, x_n) = x_{s+1}^{\,2} + \ldots + x_n^{\,2} .$$

We obtain Tamhankar's result when $X_1, \ldots, X_n$ are assumed to be mutually independent and identically distributed, and Flusser's result by assuming only $(X_1, \ldots, X_s)$ and $(X_{s+1}, \ldots, X_n)$ are independent $1 \le s < n$.

Example 2.1 which follows shows that the independence break of the transformed variables is free to occur anywhere between 1 and r. Examples 2.2 and 2.3 with the additional assumption that $X_0, \ldots, X_n$ are mutually independent are the subject of the paper by Kotlarski [6].

<u>Example 2.1</u>: Let $\vec{W} = (X_1,X_2)$ and $\vec{Z} = (X_3,X_4)$ be independent with continuous positive density functions f_{12} and f_{34}. Let a_j, b_j, $j = 1, 2, 3, 4$, be real numbers such that

$$k_1 = \frac{1}{a_1a_4 - a_2a_3} ,$$

$$k_2 = \frac{1}{b_1b_4 - b_2b_3}$$

are finite and positive.

Let

$$\lim_{x_1, x_2 \to 0} f(x_1, x_2) = (2\pi k_1)^{-1} \quad ; \quad \lim_{x_3, x_4 \to 0} f(x_3, x_4) = (2\pi k_2)^{-1} .$$

Transform as follows

$$x_1 = k_1 \; r \; (a_4 \cos \theta_1 - a_2 \sin \theta_1) \cos \theta_2 \cos \theta_3$$

$$x_2 = k_1 \; r \; (a_1 \sin \theta_1 - a_3 \cos \theta_1) \cos \theta_2 \cos \theta_3$$

$$x_3 = k_2 \; r \; (b_4 \sin \theta_2 \cos \theta_3 - b_2 \sin \theta_3)$$

$$x_4 = k_2 \; r \; (b_1 \sin \theta_3 - b_3 \sin \theta_2 \cos \theta_3)$$

where

$$x_j \in R, \; j = 1, \ldots, 4 .$$

$$0 \leq r \leq \infty$$

$$0 \leq \theta_1 < 2\pi$$

$$-\pi/2 \leq \theta_j \leq \pi/2$$

$$j = 1,2 .$$

Put

$$\Psi(r) = r^2$$

$$\Psi_1(x_1, x_2) = c_1 x_1^2 + c_2 x_1 x_2 + c_3 x_2^2 \qquad 0 \leq r < \infty$$

$$\Psi_2(x_3, x_4) = c_4 x_3^2 + c_5 x_3 x_4 + c_6 x_4^2 \qquad x_j \in R, \quad j = 1, \ldots, 4 .$$

where the c_j, $j = 1, \ldots, 6$, are appropriate combinations of the a_j and b_j so that the transformation leads to the equation

$$\Psi(r) = \Psi_1(x_1, x_2) + \Psi_2(x_3, x_4), \quad 0 \leq r < \infty, \quad x_j \in R, \quad j = 1, \ldots, 4 .$$

The necessary and sufficient condition that (X_1,X_2) and (X_3,X_4) be distributed according to

$$f_{12}(x_1,x_2) = \frac{1}{2\pi k_1} e^{-\frac{1}{2}\Psi_1(x_1,x_2)},$$

$$x_j \in R, \quad j = 1, \ldots, 4 .$$

$$f_{34}(x_3,x_4) = \frac{1}{2\pi k_2} e^{-\frac{1}{2}\Psi_2(x_3,x_4)},$$

is that R and $(\theta_1, \theta_2, \theta_3)$ or (R,θ_1) and (θ_2,θ_3) be independent.

Recall the gamma distribution $G(p,a)$ with parameters $p > 0$, $a > 0$, which is given by the density

$$f(x) = \begin{cases} \dfrac{a^p x^{p-1} e^{-ax}}{\Gamma(p)}, & x > 0 \\ 0 & \text{otherwise} . \end{cases}$$

<u>Example 2.2</u>: Let $\vec{W} = (X_0, \ldots, X_s)$ and $\vec{Z} = (X_{s+1}, \ldots, X_n)$, $(0 \le s < n)$, be two independent random vectors with positive coordinates. Assume that the probability density functions f and g of $\vec{W}$ and $\vec{Z}$ are continuous for positive arguments and that there exist limits

$$\lim_{x_0, \ldots, x_s \to 0} \frac{f(x_0, \ldots, x_s)}{x_0^{p_0-1} \cdots x_s^{p_n-1}} = A_2 > 0$$

$$\lim_{x_{s+1}, \ldots, x_n \to 0} \frac{g(x_{s+1}, \ldots, x_n)}{x_{s+1}^{p_{s+1}-1} \cdots x_n^{p_n-1}} = A_2 > 0$$

for some set of positive p_i, $i = 0, \ldots, n$.

Transform as follows:

$$\begin{aligned} X_0 &= Y\,Y_1\,Y_2 \;\cdot\;\cdot\;\cdot\; Y_n \\ X_1 &= Y\,(1-Y_1)\,Y_2 \cdots Y_n \\ X_2 &= Y\,(1-Y_2) \cdots Y_n \qquad y > 0, \\ &\vdots \qquad\qquad\qquad 0 < y_i < 1, \quad i = 1, \ldots, n \\ X_{n-1} &= Y\,(1-Y_{n-1})\,Y_n \\ X_n &= Y\,(1-Y_n)\,. \end{aligned}$$

The necessary and sufficient condition for all X_i to be mutually independent and distributed $G(p_i, a)$ is that there exist an integer q, $0 \leq q < n$, such that $(Y, \ldots, Y_q)$ and $(Y_{q+1}, \ldots, Y_n)$ are independent.

<u>Example 2.3</u>: Assume $\vec{W}$ and $\vec{Z}$ are as in Example 2.2. Transform as follows:

$$\begin{aligned} X_0 &= YY_1 \qquad\qquad y > 0,\ (y_1, \ldots, y_n) \in Q \\ X_1 &= YY_2 \\ &\vdots \qquad\qquad Q = \{(y_1, \ldots, y_n) \,|\, \textstyle\sum_{i=1}^{n} y_i = 1; \\ X_{n-1} &= YY_n \qquad\qquad y_i \geq 0,\ i = 1, \ldots, n\}\ . \\ X_n &= Y(1 - Y_1 - \ldots - Y_n)\ . \end{aligned}$$

The necessary and sufficient condition for all X_i to be mutually independent and distributed $G(p_i, a)$ is that Y and $(Y_i, \ldots, Y_n)$ be independent.

CHAPTER III

GENERALIZATION OF THE THEOREM OF KAGAN AND SHALAEVSKI

A restatement of the theorem of Kagan and Shalaevski is included for reference.

Theorem 3.1: Let $X_1, \ldots, X_n$ be i.i.d. and suppose $Y = \sum_{j=1}^{n} (X_j + a_j)^2$, $a_j \in R$, has a distribution which depends on the a_j only through $\sum_{j=1}^{n} a_j^2$. Then the common distribution of the X_j is $N(0, \sigma)$.

The original method of proof as given by the authors was to twice differentiate the function $h(a_1, \ldots, a_n) = E \exp[-\sum_{j=1}^{n} (X_i + a_i)^2]$. When the first generalization was made this method was abandoned and replaced by the solution of the Cauchy equation. We include the proof of this theorem as originally done to give contrast to the proof of the most recent result.

Theorem 3.2: Let $(X_1, \ldots, X_m)$ and $(X_{m+1}, \ldots, X_n)$, $1 \leq m < n$, be independent and let $Y = \sum_{j=1}^{n} (X_j + a_j)^2$ have a distribution which depends only on $\sum_{j=1}^{n} a_j^2$, $a_j \in R$. Then all X_j are independent and distributed normally with zero means and common variance.

Proof: Let $F(x_1, \ldots, x_n)$, $F_0(x_1, \ldots, x_m)$, $F_1(x_{m+1}, \ldots, x_n)$ be the distribution functions for the vectors $(X_1, \ldots, X_n)$,

$(X_1, \ldots, X_m)$, and $(X_{m+1}, \ldots, X_n)$, respectively. Define a function h on R^n as follows:

$$h(a_1, \ldots, a_n) = E \exp\left(-\sum_{j=1}^{n} (X_j + a_j)^2\right)$$

$$= \int_{R^n} \exp\left(-\sum_{j=1}^{n} (x_j + a_j)^2\right) dF(x_1, \ldots, x_n), \quad a_j \in R . \tag{3.2.1}$$

Note that h is continuous with respect to $a_1, \ldots, a_n$ by the convergence theorem for Lebesgue integrals (see Loeve [8, p. 125]). Set

$$G(u_1, \ldots, u_n) = \int_{-\infty}^{u_1} \cdots \int_{-\infty}^{u_n} \exp\left(-\sum_{j=1}^{n} x_j^2\right) dF(x_1, \ldots, x_n), \quad u_j \in R . \tag{3.2.2}$$

Because of the independence of $(X_1, \ldots, X_m)$ and $(X_{m+1}, \ldots, X_n)$ we may write

$$G_0(u_1, \ldots, u_m) G_1(u_{m+1}, \ldots, u_n)$$

$$= \int_{-\infty}^{u_1} \cdots \int_{-\infty}^{u_m} \exp\left(-\sum_{j=1}^{m} x_j^2\right) dF_0(x_1, \ldots, x_m) \tag{3.2.3}$$

$$\cdot \int_{-\infty}^{u_{m+1}} \cdots \int_{-\infty}^{u_n} \exp\left(-\sum_{j=m+1}^{n} x_j^2\right) dF_1(x_{m+1}, \ldots, x_n), \quad u_j \in R .$$

Now by expanding the exponent in (3.2.1) and using (3.2.2) and (3.2.3) we may write

$$h(a_1, \ldots, a_n) = \int_{R^m} \exp\left(-\sum_{j=1}^{m} 2a_j x_j\right) \exp\left(-\sum_{j=1}^{m} a_j^2\right) dG_0(x_1, \ldots, x_m)$$

$$\cdot \int_{R^{n-m}} \exp\left(-\sum_{j=m+1}^{n} 2a_j x_j\right) \exp\left(-\sum_{j=m+1}^{n} a_j^2\right) dG_1(x_{m+1}, \ldots, x_n),$$

$$a_j \in R . \tag{3.2.4}$$

But by hypothesis

$$h(a_1, \dots, a_n) = \Psi(\sum_{j=1}^{n} a_j^2), \qquad a_j \in R \; ; \tag{3.2.5}$$

hence combining (3.2.4) and (3.2.5)

$$\int_{R^m} \exp(-\sum_{j=1}^{m} 2a_jx_j) \exp(-\sum_{j=1}^{m} a_j^2) \, dG_0(x_1, \dots, x_m)$$

$$\cdot \int_{R^{n-m}} \exp(-\sum_{j=m+1}^{n} 2a_jx_j) \exp(-\sum_{j=m+1}^{n} a_j^2) \, dG_1(x_{n+1}, \dots, x_n)$$

$$= \Psi(\sum_{j=1}^{n} a_j^2), \qquad a_j \in R \, . \tag{3.2.6}$$

Setting

$$\Psi_1(t) = \Psi(t) \cdot e^t \; , \qquad t \geq 0 \; , \tag{3.2.7}$$

we see that 3.2.6 becomes

$$\int_{R^m} \exp(-\sum_{j=1}^{m} 2a_jx_j) \, dG_0(x_1, \dots, x_m) \int_{R^{n-m}} \exp(-\sum_{j=m+1}^{n} 2a_jx_j) dG_1(x_{m+1}, \dots, x_n)$$

$$= \Psi_1(\sum_{j=1}^{n} a_j^2) \; , \qquad a_j \in R \, . \tag{3.2.8}$$

Let

$$b_0 = \int_{R^m} dG_0(x_1, \dots, x_m)$$

$$b_1 = \int_{R^{n-m}} dG_1(x_{m+1}, \dots, x_n) \tag{3.2.9}$$

and note that neither b_0 nor b_1 is zero. Evaluating (3.2.8) first when $a_{m+1} = a_{m+2} = \dots = a_n = 0$ and then when $a_1 = a_2 = \dots = a_m = 0$ we find

$$\Psi_1(\sum_{j=1}^{n} a_j^2) = b_1 \int_{R^m} \exp(-\sum_{j=1}^{m} 2a_j x_j) dG_0(x_1, \ldots, x_m) , \qquad a_j \in R$$

$$\Psi_1(\sum_{j=m+1}^{n} a_j^2) = b_0 \int_{R^{n-m}} \exp(-\sum_{j=m+1}^{n} 2a_j x_j) dG_1(x_{m+1}, \ldots, x_n), \quad a_j \in R . \tag{3.2.10}$$

Substituting (3.2.10) into (3.2.8)

$$\frac{\Psi_1(\sum_{j=1}^{m} a_j^2)\Psi_1(\sum_{j=m+1}^{n} a_i^2)}{b_0 b_1} = \Psi_1(\sum_{j=1}^{n} a_j^2), \quad a_j \in R. \tag{3.2.11}$$

Dividing both sides of (3.2.11) by $b_0 b_1$ and setting

$$\frac{\Psi_1(t)}{b_0 b_1} = \Psi_2(t), \qquad t \geq 0 \tag{3.2.12}$$

we have

$$\Psi_2(\sum_{j=1}^{m} a_j^2)\Psi_2(\sum_{j=m+1}^{n} a_j^2) = \Psi_2(\sum_{j=1}^{n} a_j^2), \quad a_j \in R \tag{3.2.13}$$

which is a form of the Cauchy equation (see Aczel [1, p. 31]) and, because the continuity of Ψ_2 follows from the continuity of h, has solution

$$\Psi_2(t) = e^{ct}, \qquad t \geq 0 \tag{3.2.14}$$

where c is a real constant. From (3.2.12),

$$\Psi_1(t) = b_0 b_1 e^{ct}, \qquad t \geq 0 . \tag{3.2.15}$$

Now from (3.2.2), (3.2.5), and (3.2.7) we may write

$$\Psi_1(\sum_{j=1}^{n} a_j^2) = \int_{R^n} \exp(-\sum_{j=1}^{n} 2a_j x_j) dG(x_1, \ldots, x_n), \quad a_j \in R; \tag{3.2.16}$$

thus from (3.2.15) and (3.2.16)

$$\int_{R^n} \exp(-\sum_{j=1}^{n} 2a_j x_j) dG(x_1, \ldots, x_n) = b_0 b_1 \exp(c \sum_{j=1}^{n} a_j^2), \quad a_j \in R . \tag{3.2.17}$$

But (3.2.17) is an n-dimensional transform with parameters $2a_j, j = 1, 2, \ldots, n$, hence $dG(x_1, \ldots, x_n)$ is determined uniquely by

$$dG(x_1, \ldots, x_n) = k_1 \exp(k_2 \sum_{j=1}^{n} x_j^2) dx_1 \ldots dx_n, \quad x_j \in R . \tag{3.2.18}$$

Then from (3.2.2)

$$dF(x_1, \ldots, x_n) = k_1 \exp(k_3 \sum_{j=1}^{n} x_j^2) dx_1 \ldots dx_n, \quad x_j \in R .$$

Since F is a distribution function k_1 and k_3 must be given by

$$k_3 = -(2\sigma^2)^{-1} ,$$

$$k_1 = [(2\pi)^{n/2} \sigma^n]^{-1} ,$$

which shows all X_j to be distributed independently and normally with zero mean and common variance.

In attempting to generalize further it was noticed that the statistic $Y = \sum_{j=1}^{n}(X_i + a_i)^2$ is essentially a function of vectors $\vec{X}$ and $\vec{a}$ and may be written as $Y = (\vec{X} + \vec{a})\, I\, (\vec{X} + \vec{a})'$ where $'$ denotes transposition and I is the $n \times n$ identity matrix. Further, the statistic Y being a sum may be freely broken; i.e., $Y = \sum_{j=1}^{m}(X_j + a_j)^2 + \sum_{j=m+1}^{n}(X_i + a_i)^2$. The following theorems and corollaries are natural extensions of Theorem 3.2; hence, some of the proofs are omitted. Theorems 3.3 and 3.4 taken together allow a characterization of the normal distribution.

Theorem 3.3: Let $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_m)$ be real random row vectors. Let

$$U = U_{n \times n} \qquad V = V_{m \times m}$$

$$A = A_{n \times n} \qquad B = B_{m \times m}$$

be real, symmetric, positive definite matrices with $A + U^{-1}$, $B + V^{-1}$ nonsingular. Denote

$$C = [A(A + U^{-1})^{-1} - I]A$$

$$D = [B(B + V^{-1})^{-1} - I]B \ .$$

For all row vectors $a \in R^n$, $b \in R^m$ define

$$Z = (X - a)A(X - a)' + (Y - b)B(Y - b)'$$

where $'$ again means transposition. If X and Y are independent and distributed $N(0,U)$ and $N(0,V)$, respectively, then the expectation of Z depends only on $aCa' + bDb'$.

Proof: Define the following functions

$$\begin{cases} \varphi_X(a) = E\exp[-\tfrac{1}{2}(X-a)A(X-a)'] & a \in R^n \\ \varphi_Y(b) = E\exp[-\tfrac{1}{2}(Y-b)B(Y-b)'] & b \in R^m \ . \\ \varphi_Z(a,b) = E\exp[-\tfrac{1}{2}(X-a)A(X-a)' - \tfrac{1}{2}(Y-b)B(Y-b)'] \end{cases} \tag{3.3.1}$$

By hypothesis $X \sim N(0,U)$, thus evaluating $\varphi_X(a)$ we find

$$\varphi_X(a) = \frac{1}{(2\pi)^{n/2}\sqrt{|U|}} \int_{R^n} \exp[-\tfrac{1}{2}(x-a)A(x-a)']\exp[-\tfrac{1}{2}xU^{-1}x'] \, dx_1 \ldots dx_n \ , \quad a \in R^n \ . \tag{3.3.2}$$

Expanding the exponent and simplifying

$$\varphi_X(a) = \frac{(2\pi)^{n/2}\sqrt{|(A+U^{-1})^{-1}|}}{(2\pi)^{n/2}\sqrt{|U|}} \exp[-\tfrac{1}{2}aAa'] \; .$$

$$\cdot \frac{1}{(2\pi)^{n/2}\sqrt{|(A+U^{-1})^{-1}|}} \int_{R^n} \exp[aAx' - \tfrac{1}{2}\, x(A+U^{-1})x']$$

$$\cdot\ dx_1 \ldots dx_n \, , \qquad a \in R^n \; . \qquad (3.3.3)$$

The final portion of (3.3.3) is the moment generating function of a normal vector with covariance matrix $(A + U^{-1})^{-1}$ evaluated at point aA (see Moran [9, p. 272]), thus

$$\varphi_X(a) = \sqrt{\frac{|(A+U^{-1})^{-1}|}{|U|}} \exp[-\tfrac{1}{2}aAa']\exp[\tfrac{1}{2}aA[(A+U^{-1})^{-1}]Aa'] , \quad a \in R^n. \qquad (3.3.4)$$

Combining exponents and recalling the definition of C

$$\varphi_X(a) = k_1 \exp[\tfrac{1}{2}aCa'] , \quad a \in R^n \; . \qquad (3.3.5)$$

Similarly

$$\varphi_Y(b) = k_2 \exp[\tfrac{1}{2}bDb'] , \quad b \in R^m \; . \qquad (3.3.6)$$

From the independence of X and Y we have that

$$\varphi_Z(a,b) = \varphi_X(a)\varphi_Y(b), \quad a \in R^n, \quad b \in R^m \; .$$

Hence from (3.3.5) and (3.3.6)

$$\varphi_Z(a,b) = k_3 \exp[\tfrac{1}{2}(aCa' + bDb'] , \quad a \in R^n, \quad b \in R^m \; ,$$

which is sufficient to show the expectation of Z is dependent only on $aCa' + bDb'$.

Corollary 3.1: Let $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_m)$ be real random row vectors. Let $U = U_{n \times n}$ and $V = V_{m \times m}$ be real, symmetric, positive definite matrices. For all $a \in R^n$, $b \in R^m$ define

$$Z = (X - a)U^{-1}(X - a)' + (Y - b)V^{-1}(Y - b)' .$$

If X and Y are independent and distributed $N(0,U)$ and $N(0,V)$ respectively then the distribution of Z depends only on $aU^{-1}a' + bV^{-1}b'$.

Theorem 3.4: Let $X = (X_1, \ldots, X_n)$, $Y = (Y_1, \ldots, Y_n)$ be two real independent random vectors. Let A, C be real $n \times n$ matrices and B, D be real $m \times m$ matrices with $A - cC$ and $B - cD$ invertable for suitable real constant c. Denote

$$P = P(c) = A'(A - cC)^{-1}A$$

$$Q = Q(c) = B'(B - cD)^{-1}B$$

and require that $P - A$ and $Q - B$ be positive definite and symmetric. For all $a \in R^n$, $b \in R^m$ define

$$Z = (X-a)A(X-a)' + (Y-b)B(Y-b)' .$$

If the distribution of Z depends only on $aCa' + bDb'$ then X and Y are distributed normally with zero means and covariance matrices $(P - A)^{-1}$ and $(Q - B)^{-1}$ respectively.

Corollary 3.2: Let X and Y be real independent random row vectors. Let $A = A_{n \times n}$ and $B = B_{m \times m}$ be real, symmetric, positive definite matrices. For all $a \in R^n$, $b \in R^m$ define

$$Z = (X - a)A(X - a)' + (Y - b)B(Y - b)' .$$

If the distribution of Z depends only on $aAa' + bBb'$ then X and Y are distributed $N(0, cA^{-1})$ and $N(0, cB^{-1})$ respectively for a suitable constant c.

Further generalization hinged on the recovery of the original distribution from the integral transform. Although tables are extensive, they lacked the exact transforms of interest. This problem was overcome by introducing convolutions of probability measures and requiring the statistic to be sufficiently close to a probability density function. We begin with some preliminary definitions and notations.

Let μ be a probability measure (p.m.) and f a p.d.f. defined on real n-space R^n. The respective Fourier transforms $\hat{\mu}$ and $\hat{f}$ will be given by

$$\hat{\mu}(t) = \int_{R^n} e^{itx} \mu(dx) , \quad t \in R^n$$

$$\hat{f}(t) = \int_{R^n} e^{itx} f(x)dx , \quad t \in R^n .$$

We define the convolution of f and μ by

$$(f * \mu)(a) = \int_{R^n} f(a - x)\mu(dx) , \quad a \in R^n$$

and note that

$$(f * \mu)(a) = E_{\mu} f(a - X) , \quad a \in R^n .$$

Here E_μ denotes the expectation under the measure μ and when no confusion results the μ will be supressed. The convolution $f * \mu$ is a p.d.f. which corresponds to the sum of two independent random vectors, one with p.d.f. f and the other with p.m. μ (see Moran [9, p. 230]) and their Fourier transforms are related by

$$\widehat{f * \mu} = \hat{f} \cdot \hat{\mu} .$$

The following conditions are notation will be used for the theorems and corollaries which follow.

I. $Y_1 = (X_1, \ldots, X_n)$ and $Y_2 = (X_{n+1}, \ldots, X_{n+m})$ are real independent random vectors with p.m.'s μ_1 and μ_2.

II. The functions f_1 and f_2 are p.d.f.'s on R^n and R^m, respectively.

III. The Fourier transforms $\hat{f}_1$ and $\hat{f}_2$ do not vanish.

IV. For $\mathcal{A}_1 \subset R^n$, $\mathcal{A}_2 \subset R^m$ and $\mathfrak{D}$ being one of $(0,\infty)$, $[1,\infty)$, or $(0,1]$, the functions g_1 and g_2 defined on R^n and R^m, respectively, satisfy

$$\left.\begin{array}{l} g_k : \mathcal{A}_k \xrightarrow{\text{onto}} \mathfrak{D} \\ \\ g_k(x) = 0 , \quad x \notin \mathcal{A}_k . \end{array}\right. \qquad k = 1,2$$

Notation: Denote

$$\eta_k(a_k) = (f_k * \mu_k)(a_k) , \quad k = 1,2 , \quad a_1 \in R^n , \quad a_2 \in R^m ,$$

and let a_k^* denote a point of $\mathcal{A}_k$ such that $g_k(a_k^*) = 1$, $k = 1, 2$.

V. $\eta_1(a_1^*) \neq 0 \neq \eta_2(a_2^*)$.

VI. For all $c \in \mathcal{C} \subset R$ and for suitable c_1 and c_2 dependent on c,

$$h_k(a_k;c) = c_k g_k^c(a_k)$$

is a p.d.f. $k=1,2$, $a_1 \in R^n$, $a_2 \in R^m$.

VII. There exists a set $\mathcal{C}' \subset \mathcal{C}$ such that for $c \in \mathcal{C}'$ the equations

$$\hat{f}_k(t)\varphi_k(t;c) = \hat{h}_k(t;c) \qquad \begin{matrix} t \in R^n & \text{if} & k=1 \\ t \in R^m & \text{if} & k=2 \end{matrix} , \tag{3.6.1}$$

have solutions for unknown φ_k which are Fourier transforms of p.m.'s on R^n and R^m, respectively.

Whenever I and II are satisfied we shall denote

$$Z(a_1,a_2) = f_1(a_1 - Y_1)f_2(a_2 - Y_2) \ , \quad a_1 \in R^n, \quad a_2 \in R^m \ . \tag{3.6.2}$$

<u>Theorem 3.5</u>: If assumptions I and II are satisfied then

$$E\ Z(a_1,a_2) = \eta_1(a_1)\eta_2(a_2) \ , \quad a_1 \in R^n, \quad a_2 \in R^m \ .$$

<u>Proof</u>: Since Y_1 and Y_2 are specified independent in I we may write for $a_1 \in R^n$, $a_2 \in R^m$,

$$E\ Z(a_1,a_2) = E\ f_1(a_1 - Y_1)f_2(a_2 - Y_2) =$$

$$= E\ f_1(a_1 - Y_1)\, E\, f_2(a_2 - Y_2) = \eta_1(a_1)\eta_2(a_2) \ .$$

<u>Theorem 3.6</u>: Suppose assumptions I - VII are satisfied and the expectation of $Z(a_1,a_2)$ depends on a_1 and a_2 only through a

function of $g_1(a_1) \cdot g_2(a_2)$ which is continuous at at least one point of $\mathfrak{D}$ and is zero at zero. Then the p.m.'s μ_1 and μ_2 are determined up to a parameter $c \in \mathcal{C}'$ by

$$\hat{\mu}_1(t;c) = \frac{\hat{h}_1(t;c)}{\hat{f}_1(t)}, \qquad t \in R^n$$
$$\hat{\mu}_2(s;c) = \frac{\hat{h}_2(s;c)}{\hat{f}_2(s)}, \qquad s \in R^m \quad . \tag{3.6.3}$$

<u>Remark</u>: These two theorems along with an additional condition which specified c may be used to characterize μ_1 and μ_2.

<u>Proof of Theorem 3.6</u>: By Theorem 1 we have

$$E\ Z(a_1,a_2) = \eta_1(a_1)\eta_2(a_2), \quad a_1 \in R^n, \quad a_2 \in R^m \ . \tag{3.6.4}$$

On the other hand by hypothesis $E\ Z(a_1,a_2)$ may be expressed as

$$E\ Z(a_1,a_2) = \rho(g_1(a_1)\cdot g_2(a_2)), \quad a_1 \in R^n, \quad a_2 \in R^m , \tag{3.6.5}$$

with ρ being continuous at at least one point of $\mathfrak{D}$ and zero at zero. Combining (3.6.4) and (3.6.5) we obtain the functional equation

$$\eta_1(a_1)\eta_2(a_2) = \rho(g_1(a_1)\cdot g_2(a_2)), \quad a_1 \in R^n, \quad a_2 \in R^m , \tag{3.6.6}$$

in which η_1 and η_2 are unknown. Evaluating (3.6.6) when $a_2 = a_2^*$ and then when $a_1 = a_1^*$ we see that because of IV

$$\eta_1(a_1)\eta_2(a_2^*) = \rho(g_1(a_1)), \qquad a_1 \in R^n$$
$$\eta_2(a_2)\eta_2(a_1^*) = \rho(g_2(a_2)), \qquad a_2 \in R^m$$

or because V requires $\eta_1(a_1^*) \neq 0 \neq \eta_2(a_2^*)$,

$$\eta_1(a_1) = \frac{\rho(g_1(a_1))}{\eta_2(a_2^*)} , \qquad a_1 \in R^n$$
$$\eta_2(a_2) = \frac{\rho(g_2(a_2))}{\eta_1(a_1^*)} , \qquad a_2 \in R^m . \tag{3.6.7}$$

Define a function γ by

$$\gamma(t) = \frac{\rho(t)}{\eta_1(a_1^*)\eta_2(a_2^*)} , \qquad t \in \mathfrak{D} \cup \{0\} . \tag{3.6.8}$$

Note that since ρ is zero at zero and continuous at some point of $\mathfrak{D}$ the same is true of γ. Also since $g_k \equiv 0$ on the complement of $\mathfrak{a}_k$, $k = 1, 2$, it follows from (3.6.7) that $\eta_k \equiv 0$ on complement of $\mathfrak{a}_k$, $k = 1, 2$. We conclude that η_1 and η_2 are unknown only on $\mathfrak{a}_1$ and $\mathfrak{a}_2$, hence restrict ourselves to those sets. Now using (3.6.7) and (3.6.8) in (3.6.6) we get

$$\gamma(g_1(a_1)\cdot g_2(a_2)) = \gamma(g_1(a_1))\gamma(g_2(a_2)) , \quad a_1 \in \mathfrak{a}_1, \quad a_2 \in \mathfrak{a}_2 , \tag{3.6.9}$$

or setting $g_1(a_1) = t_1$, $g_2(a_2) = t_2$,

$$\gamma(t_1 \cdot t_2) = \gamma(t_1)\gamma(t_2) , \quad t_1, t_2 \in \mathfrak{D} . \tag{3.6.10}$$

This is the Cauchy equation (see Aczel [1, p. 38]) which has most general solutions for γ continuous at at least one point

$$\gamma(t) = t^c \qquad \text{or} \qquad \gamma(t) \equiv 0 , \quad t \in \mathfrak{D} , \tag{3.6.11}$$

and c a real constant. We now retrace our steps. From (3.6.8)

$$\rho(t) = \eta_1(a_1^*)\eta_2(a_2^*)t^c \quad \text{or} \quad \rho(t) \equiv 0 , \quad t \in \mathfrak{D} . \tag{3.6.12}$$

Putting (3.6.12) into (3.6.7) and recalling that $g_k \equiv 0$ outside $\mathfrak{a}_k$, k = 1, 2, we obtain

$$\begin{cases} \eta_1(a_1) = \eta_1(a_1^*)g_1^c(a_1)\ , & a_1 \in R^n \\ \eta_2(a_2) = \eta_2(a_2^*)g_2^c(a_2)\ , & a_2 \in R^m\ , \end{cases} \tag{3.6.13}$$

or

$$\eta_1(a_1) \equiv 0 \equiv \eta_2(a_2)\ , \quad a_1 \in R^n, \quad a_2 \in R^m\ .$$

However, we see from the requirement $\eta_1(a_1^*) \neq 0 \neq \eta_2(a_2^*)$ that the latter case is impossible. Recalling the definition of η_k, (3.6.13) becomes

$$\begin{cases} (f_1 * \mu_1)(a_1) = \eta_1(a_1^*)g_1^c(a_1), & a_1 \in R^n \\ (f_2 * \mu_2)(a_2) = \eta_2(a_2^*)g_2^c(a_2), & a_2 \in R^m\ . \end{cases} \tag{3.6.14}$$

The left sides of (3.6.14) are p.d.f.'s, as noted when the convolution was defined. The right sides are also p.d.f.'s for $c \in \mathbb{C}$ as given in VI. Thus, for $c \in \mathbb{C}$ (3.6.14) becomes

$$\begin{cases} (f_1 * \mu_1)(a_1) = h_1(a_1;c)\ , & a_1 \in R^n \\ (f_2 * \mu_2)(a_2) = h_2(a_2;c)\ , & a_2 \in R^m\ . \end{cases}$$

Since μ_k is a p.m., $k = 1, 2$, we have our second restriction on the parameter c and may state that for $c \in \mathbb{C}'$ as given in VII the unknown μ_k are determined by

$$\begin{cases} \hat{f}_1(t) \cdot \hat{\mu}_1(t;c) = \hat{h}_1(t;c), & t \in R^n \\ \hat{f}_2(s) \cdot \hat{\mu}_2(s;c) = \hat{h}_2(s;c), & s \in R^m\ . \end{cases}$$

Finally, since III requires $\hat{f}_1$ and $\hat{f}_2$ to be nonzero, we see the above equations are equivalent to those given in (3.6.3) which completes the proof.

Remarks: (i) With minor alterations the functions f_1 and f_2 may be replaced by constant multiples of p.d.f.'s.

(ii) By setting $\Psi_k(a_k) = \log g_k(a_k)$, $a_k \in \mathcal{A}_k$, $k = 1, 2$, the dependence on the product $g_1(a_1)\cdot g_2(a_2)$ becomes a dependence on the sum $\Psi_1(a_1) + \Psi_2(a_2)$, $a_1 \in \mathcal{A}_1$, $a_2 \in \mathcal{A}_2$.

The following examples and corollaries illustrate the application of Theorems 3.5 and 3.6.

Example 3.1: Assume $\alpha > 0$, $0 < p_k < 1$, $p_k + q_k = 1$, $k = 1, 2, \ldots, m+n$. Assume $Y_1 = X_1, \ldots, X_n)$ and $Y_2 = (X_{n=1}, \ldots, X_{n+m})$ are real independent random vectors. Denote

$$f_1(x_1, \ldots, x_n) = \begin{cases} \prod_{j=1}^{n} \frac{\alpha^{p_j}}{\Gamma(p_j)} x_j^{p_j - 1} e^{-\alpha x_j} & \text{all } x_j > 0 \\ 0 & \text{elsewhere} \end{cases}$$

$$f_2(x_{n+1}, \ldots, x_{n+m}) = \begin{cases} \prod_{j=n+1}^{n+m} \frac{\alpha^{p_j}}{\Gamma(p_j)} x_j^{p_j - 1} e^{-\alpha x_j} & \text{all } x_j > 0 \\ 0 & \text{elsewhere} \end{cases}$$

and

$$g_1(a_1, \ldots, a_n) = \begin{cases} \exp\left[-\sum_{j=1}^{n} a_j\right] & \text{all } a_j \geq 0 \\ 0 & \text{elsewhere} \end{cases}$$

$$g_2(a_{n+1}, \ldots, a_{n+m}) = \begin{cases} \exp\left[-\sum_{j=n+1}^{n+m} a_j\right] & \text{all } a_j \geq 0 \\ 0 & \text{elsewhere} . \end{cases}$$

Assume for some $1 \leq k \leq n+m$ that $E X_k = \frac{q_k}{\alpha}$.

Corollary 3.3: The expectation of

$$Z(b_1,b_2) = f_1(b_1 - Y_1)f_2(b_2 - Y_2) \qquad \begin{array}{l} b_1 = (a_1, \ldots, a_n) \in R^n \\ b_2 = (a_{n+1}, \ldots, a_{n+m}) \in R^m \end{array}$$

depends on b_1 and b_2 only through a continuous function of

$$\sum_{j=1}^{n} a_j + \sum_{j=n+1}^{n+m} a_j$$

if and only if all coordinates of Y_1 and Y_2 are mutually independent with $X_j \sim G(q_j,\alpha)$.

Proof: It is evident that Theorem 3.5 applies. Assumptions I - VI are satisfied for $\mathcal{C} = (0,\infty)$ and

$$h_1(b_1;c) = \begin{cases} c^{-n} \exp\left[-c \sum_{j=1}^{n} a_j\right] & \text{all } a_j \geq 0 \\ 0 & \text{elsewhere} \end{cases}$$

$$h_2(b_2;c) = \begin{cases} c^{-m} \exp\left[-c \sum_{j=n+1}^{n+m} a_j\right] & \text{all } a_j \geq 0 \\ 0 & \text{elsewhere} \end{cases}$$

The corresponding characteristic functions are

$$\hat{h}_1(t;c) = \prod_{j=1}^{n} \left(1 - \frac{it_j}{c}\right)^{-1} , \qquad t_j \in R$$

$$\hat{h}_2(s;c) = \prod_{j=n+1}^{n+m} \left(1 - \frac{is_j}{\alpha}\right)^{-1} , \qquad s_j \in R .$$

We see that the equations

$$\varphi_1(t;c) = \prod_{j=1}^{n} \frac{\left(1 - \frac{it_j}{c}\right)^{-1}}{\left(1 - \frac{is_j}{\alpha}\right)^{-p_j}} , \qquad s_j \in R$$

have solutions for φ_k which are Fourier transforms of p.m.'s at least for $c = \alpha$. Thus VII is satisfied for some set $\mathcal{C}' \subset \mathcal{C}$, $\mathcal{C}'$ containing at least α. Thus Theorem 3.6 applies. From (3.6.15) we see that the characteristic function corresponding to X_j is

$$\hat{F}_j(t) = \left(1 - \frac{it}{c}\right)^{-1} \left(1 - \frac{it}{\alpha}\right)^{p_j} , \quad t \in R ,$$

where F_j denotes the distribution function of X_j. Then from the relationship $-i\hat{F}_j'(0) = E X_j$ ($'$ denotes derivative here) we see that the assumption $E X_k = \frac{q_k}{\alpha}$ actually forces $c = \alpha$. Hence the characteristic functions are

$$\hat{F}_j(t) = \left(1 - \frac{it}{\alpha}\right)^{-q_j} , \qquad j = 1, 2, \ldots, n+m ,$$

and each $X_j \sim G(q_k, \alpha)$.

Example 3.2: The previous example characterized the gamma distribution under the strict assumption that $p_k + q_k = 1$. This example shows that the assumption may be avoided and q_k allowed more freedom. For simplicity we take $m = n = 1$. Assume α, p_k, q_k are positive and $p_k + q_k \geq 1$, $k = 1, 2$. Assume X_1 and X_2 are real independent random variables and $E X_1 = \frac{q_1}{\alpha}$. Denote

$$f_k(x) = \begin{cases} \dfrac{\alpha^{p_k}}{\Gamma(p_k)} x^{p_k - 1} e^{-\alpha x} & x > 0 \\ 0 & x \leq 0 \end{cases} \qquad k = 1, 2 ,$$

and

$$g_k(a) = \begin{cases} \left(\dfrac{\alpha e}{p_k + q_k - 1}\right)^{p_k + q_k - 1} a^{p_k + q_k - 1} e^{-\alpha a} & a > 0 \\ 0 & a \leq 0 \end{cases} \qquad k = 1, 2 .$$

<u>Corollary 3.4</u>: The expectation of

$$Z(a_1, a_2) = f_1(a_1 - X_1) f_2(a_2 - X_2) , \qquad a_1, a_2 \in R$$

depends on a_1 and a_2 only through a continuous function of $g_1(a_1) \cdot g_2(a_2)$ if and only if $X_k \sim G(q_k, \alpha)$, $k = 1, 2$.

<u>Proof</u>: Similar to the proof of Corollary 3.3.

<u>Example 3.3</u>: Let A, C be real $n \times n$ matrices and B, D be real $m \times m$ matrices which are invertable, positive definite and symmetric. Assume $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_m)$ are real and independent. Denote

$$f_1(x) = \exp\left[-\tfrac{1}{2} x A^{-1} x'\right] , \qquad x \in R^n$$

$$f_2(y) = \exp\left[-\tfrac{1}{2} y B^{-1} y'\right] , \qquad y \in R^m$$

where $'$ denotes transposition. Also set

$$g_1(a_1) = \exp\,[-\tfrac{1}{2}\, a_1\, C^{-1}\, a_1'] \,, \quad a_1 \in R^n$$

$$g_2(a_2) = \exp\,[-\tfrac{1}{2}\, a_2\, D^{-1}\, a_2'] \,, \quad a_2 \in R^m$$

and assume there exists a non-empty set $\mathcal{C}' \subset (0,\infty)$ such that for $c \in \mathcal{C}'$, $cC - A$ and $cD - B$ are invertable and positive definite.

Corollary 3.5: The expectation of

$$Z(a_1,a_2) = \exp[-\tfrac{1}{2}(a_1 - X)A^{-1}(a_2 - X)' - \tfrac{1}{2}(a_2 - Y)B^{-1}(a_2 - Y)'] \,,$$

$$a_1 \in R^n, \qquad a_2 \in R^m$$

depends on a_1 and a_2 only through a continuous function of $a_1C^{-1}\, a_1' + a_2D^{-1}\, a_2'$ if and only if $X \sim N(0, cC - A)$ and $Y \sim N(0, cD - B)$.

Proof: We have the assumptions I - VI satisfied with $\mathcal{C} = (0,\infty)$. The Equations (3.6.1) are then equivalent to

$$\varphi_1(t;c) = \frac{\exp[-\tfrac{1}{2} t\; cCt']}{\exp[-\tfrac{1}{2} t\, At']} \,, \qquad t \in R^n$$

$$\varphi_2(s;c) = \frac{\exp[-\tfrac{1}{2} s\; cDs']}{\exp[-\tfrac{1}{2} s\, Bs']} \,, \qquad s \in R^m \,.$$

These equations have solutions which are the Fourier transforms of p.m.'s when $c \in \mathcal{C}'$ (as in example 3.3), which establishes the corollary.

Example 3.4: Assume $Y_1 = (X_1, \ldots, X_n)$ and $Y_2 = (X_{n+1}, \ldots, X_{n+m})$ are real and independent. Denote

$$g_1(x_1, \ldots, x_n) = f_1(x_1, \ldots, x_n) \exp\left[-\tfrac{1}{2} \sum_{j=1}^{n} x_j^2\right], \qquad x_j \in R$$

$$g_2(x_{n+1}, \ldots, x_{n+m}) = f_2(x_{n+1}, \ldots, x_{n+m}) = \exp\left[-\tfrac{1}{2} \sum_{j=n+1}^{n+m} x_j^2\right], \quad x_m \in R \; .$$

Take $\mathcal{C} = (0,\infty)$ and $\mathcal{C}' = (1,\infty)$.

Corollary 3.6: The expectation of

$$Z(b_1,b_2) = \exp\left[-\tfrac{1}{2} \sum_{j=1}^{n+m} (a_j - X_j)^2\right] \qquad \begin{array}{l} b_1 = (a_1, \ldots, a_n) \in R^n \\ b_2 = (a_{n+1}, \ldots, a_{n+m}) \in R^m \end{array}$$

depends on b_1 and b_2 only through a continuous function of

$$\sum_{j=1}^{n} a_j^2 + \sum_{j=n+1}^{n+m} a_j^2$$

if and only if all components of Y_1 and Y_2 are mutually independent and distributed $N(0,\sigma)$.

Proof: This is a particular case of Corollary 3.5 with all matrices taken as the identity.

Remark: Corollaries 3.5 and 3.6 are restatements of Theorem 3.3 and Theorem 3.4. If the assumptions in Corollary 3.6 are changed to require all X_j be i.i.d., then the same conclusions follow which is the original result of Kagan and Shalaevski.

CHAPTER IV

SUMMARY AND CONCLUSIONS

The purpose of this work has been to offer generalizations of two know characterization theorems for probability distributions.

The generalization of Tamhankar's [10] theorem as given in Chapter II appears to be complete and allows characterizations of the normal, gamma, and Dirichlet distributions based on the independence of transformed variables.

The work presented in Chapter III extends the result of Kagan and Shalaevski [5] and allows characterization of two useful distributions, the normal and the gamma; hence, also the Chi-square and the exponential. These characterizations are based primarily on statistic properties, in particular the fashion in which the expectation of a random variable depends on a parameter. It is almost certain that these generalizations can be extended. For example, it appears possible that the random variables may be allowed to take values in a Hilbert or Banach space, where the normal distribution is given by its characteristic function. The author intends to extend these theorems to these more abstract spaces and additionally to endeavor to change the group operators of $+$ and $-$ to more general operations. This would then allow characterizations of distributions which stem from quotients and products, as an example the Cauchy distribution.

A SELECTED BIBLIOGRAPHY

[1] Aczel, J. Lectures on Functional Equations and Their Applications, Academic Press, New York (1966).

[2] Bernstein, S. N. "On Some Property Which Characterizes the Gauss' Distribution," (Russian), Trudy Leningrad Polytechnic Inst., Vol. 3, 21 (1941).

[3] Cramer, H. "Uber eine Eigenschaft der normalen Verteilungsfunktion," Math. Zeit., Vol. 41 (1936).

[4] Flusser, P. "A Generalization of a Theorem by M. V. Tamhankar," Journal of Multivariate Analysis, Vol. 1 (September, 1971).

[5] Kagan, A., Shalaevski, O. "Characterization of Normal Law by a Property of the Non-Central χ^2-Distribution," Lietuvian Journal of Mathematics, Vol. 7, No. 1 (1967), (Russian).

[6] Kotlarski, I. I. "On Characterization of Probability Distributions by Means of Independent Statistics," Annali di Matematica pura ed applicata, (IV), Vol. 83, pp. 253-260.

[7] Kotlarski, I. I. "Una caratterizzazione della distribuzione gamma per mezzo di statistiche indipendenti," Rendiconti di Matematica, (3-4), Vol. 2, Serie VI (1969), pp. 671-675.

[8] Loeve, M. Probability Theory, D. Van Nostrand Company, Inc., New Jersey (1963).

[9] Moran, P. An Introduction to Probability Theory, Oxford University Press, London (1968).

[10] Tamhankar, M. V. "A Characterization of Normality," The Annals of Mathematical Statistics, 38, No. 6 (December, 1967), pp. 1924-1927.

VITA

W. Lyle Cook

Candidate for the Degree of

Doctor of Philosophy

Thesis: SOME CHARACTERIZATION PROBLEMS IN PROBABILITY THEORY

Major Field: Mathematics

Biographical:

Personal Data: Born in Waverly, New York, August 4, 1941, the son of Dr. and Mrs. Norman Cook.

Education: Elementary education received in Bay Shore, New York: secondary education received in Cape Girardeau, Missouri; graduated from Central High School, Cape Girardeau, Missouri, in 1958. Received the Bachelor of Science degree in Education from Southeast Missouri State College, Cape Girardeau, Missouri, in 1965. Attended Georgia State University, Atlanta, Georgia, 1967-68, and completed requirements for the Doctor of Philosophy degree at Oklahoma State University, Stillwater, Oklahoma, in July, 1972.

Professional Experience: Mathematics instructor, Hackley School, Tarrytown, New York, 1965-66, and at Pace Academy, Atlanta, Georgia, 1966-68. Graduate Assistant at Georgia State University, 1968-69, and at Oklahoma State University, 1969-72.