REDUCTION OF A CERTAIN CLASS OF COMPOSITE STATISTICAL HYPOTHESES

By George W. Brown

1. Introduction. A situation frequently met in sampling theory is the following: x has distribution $f(x, \theta)$, where θ is an unknown parameter, and for samples $(x_1, \cdots, x_n)$ there exists in the sample space E_n a family of $(n - 1)$-dimensional manifolds upon each of which the distribution is independent of θ; in addition there is a residual one-dimensional manifold available for estimating θ. For example, suppose there exists a sufficient statistic T for θ, then on the manifolds $T = T_0$ there is defined an induced distribution which is independent of the parameter.

A similar situation is observed when θ is a "location" or "scale" parameter. Let x have the distribution $f(x - a)$ for some a, then the set $(x_2 - x_1, x_3 - x_1, \cdots, x_n - x_1)$, or any equivalent set, such as $(x_2 - \bar{x}, \cdots, x_n - \bar{x})$, have a joint distribution independent of a, and there is a residual distribution corresponding to each particular configuration $(x_2 - x_1, \cdots, x_n - x_1)$. Fisher [1] and Pitman [5] have examined the residual distributions in connection with the problem of estimating scale and location parameters. In this paper we shall be concerned primarily, not with the residual distribution, but with the remainder of the sample information, corresponding to the $(n - 1)$-dimensional distribution which is independent of the parameter. It is found, in a rather broad class of distributions, that the part of the sample *not* used for estimation *determines, except for the parameter value, the original functional form of the distribution of x.*

This paper is devoted mainly to a study of particular classes of distributions having the property mentioned above. We consider also the theoretical application of this property to certain types of *composite* hypotheses which may be reduced thereby to equivalent *simple* hypotheses.[1] The principal results of this nature may be summed up as follows: If x has distribution of the form $f(x, \theta)$, where θ is either a location or scale parameter, or a vector denoting both, then there exists, in samples $(x_1, \cdots, x_n)$ a set of functions $y_i(x_1, \cdots, x_n)$, $i = 1, 2, \cdots, p$, $p < n$, having joint distribution $D(y_1, \cdots, y_p)$ independent of θ, and such that the converse statement holds, namely, *if $\{y_i\}$ have the distribution $D(y_1, \cdots, y_p)$, then x has, for some θ, a distribution of the form $f(x, \theta)$.* There is a corresponding statement when x has a distribution of the form $f(x - \Sigma a_i u_i)$, where the $\{a_i\}$ are parameters, and the $\{u_i\}$ are regression variables.

[1] We use the terms simple and composite hypotheses in the sense of Neyman and Pearson [2].

2. Location and Scale. This section is devoted to the study of functions of the sample observations which are such that their distributions determine the distribution of x, except possibly for location and scale.

It will be assumed that associated with x there is a function $F(x)$ such that

(a) $F(x)$ is monotone non-decreasing,

(b) $F(-\infty) = \lim_{x\to-\infty} F(x) = 0$, and (c) $F(\infty) = \lim_{x\to\infty} F(x) = 1$

with the normalization $F(x)$ upper semi-continuous. $F(x)$ is the probability that the random variate takes a value less than or equal to x. If $F(x)$ is associated with the random variate x we say that x has the distribution $F(x)$. If $g(x)$ is a Borel-measurable function, the Lebesgue-Stieltjes integral $\int_{-\infty}^{\infty} g(x)\, dF(x)$ is denoted by $E[g(x)]$.

The *characteristic function* $\varphi(t) = E(e^{itx})$ determines $F(x)$, that is, if $\int_{-\infty}^{\infty} e^{itx}\, dG(x) = \int_{-\infty}^{\infty} e^{itx}\, dF(x)$, then $F(x) = G(x)$.

Similarly, let $F(x_1, \cdots, x_k)$ be such that

(a) $F(x_1, \cdots, x_{i-1}, x_i + h, x_{i+1}, \cdots, x_k) \geq F(x_1, \cdots, x_i, \cdots, x_k)$ for $h > 0$ and $i = 1, 2, \cdots, k$;

(b) $\lim_{x_i\to-\infty} F(x_1, \cdots, x_k) = 0, i = 1, 2, \cdots, k$;

(c) $\lim_{x_1,\cdots,x_k\to\infty} F(x_1, \cdots, x_k) = 1$;

with the normalization $F(x_1, \cdots, x_k)$ continuous on the right in each x_i. If $F(x_1, \cdots, x_k)$ is associated with $x_1, \cdots, x_k$ we say that $x_1, \cdots, x_k$ have the joint distribution $F(x_1, \cdots, x_k)$. As before, $E[H(x_1, \cdots, x_k)] = \int_{R_k} H\, dF$, where R_k is the Euclidean k-space. It is well known that under such conditions, given Borel-measurable functions $y_i(x_1, \cdots, x_k)$, $i = 1, \cdots, p$, $p \leq k$, then $G(y_1, \cdots, y_p) = \int_{R(y)} dF(x_1, \cdots, x_k)$, where $R(y)$ is the region $[y_1(x_1, \cdots, x_k) \leq y_1, \cdots, y_p(x_1, \cdots, x_k) \leq y_p]$, is again a distribution function satisfying the conditions above. Moreover, $\int_R g(y_1, \cdots, y_p)\, dG(y_1, \cdots, y_p) = \int_{R'} g[y_1(x_1, \cdots, x_k), \cdots, y_p(x_1, \cdots, x_k)]\, dF$, where R' is the set of all points $(x_1, \cdots, x_k)$ such that $[y_1(x_1, \cdots, x_k), \cdots, y_p(x_1, \cdots, x_k)] \in R$.

If x has distribution $F(x)$, then, by definition, the set $(x_1, \cdots, x_n)$ is a sample from this distribution if $x_1, \cdots, x_n$ have the joint distribution $F(x_1) \cdots F(x_n)$.

The following theorem states that two distributions giving rise, in sampling, to the same distribution of the set $x_1 - x_n, x_2 - x_n, \cdots, x_{n-1} - x_n$, with $n \geq 3$, can differ at most by a translation, that is, the distribution of that set determines the original distribution except for location.

THEOREM IA: *Let x have the distribution $F(x)$. Denote by S the set of zeros of*

$\int e^{itx}\,dF(x)$ and denote by ϵ the g.l.b. *of $|t|$ for t in S. Suppose that the complement of S is ϵ-connected.*[2] *Suppose that x' has distribution $G(x')$, and let $x_1, \cdots, x_n$ and $x_1', \cdots, x_n'$ be samples. Then the set $w_\alpha = x_\alpha - x_n$, $\alpha = 1, \cdots, n-1$, have the same joint distribution as the set $w_\alpha' = x_\alpha' - x_n'$ if and only if there exists a constant a such that $x' + a$ and x have the same distribution.*

PROOF: The sufficiency of the condition follows immediately, since $w_\alpha' = x_\alpha' - x_n' = (x_\alpha' + a) - (x_n' + a)$.

In establishing necessity, only the fact that w_1, w_2 have the same joint distribution as w_1', w_2' is needed. This hypothesis implies that

$$E\{e^{i[t_1w_1+t_2w_2]}\} = E\{e^{i[t_1w_1'+t_2w_2']}\},$$

that is,

$$E\{e^{i[t_1(x_1-x_n)+t_2(x_2-x_n)]}\} = E\{e^{i[t_1(x_1'-x_n')+t_2(x_2'-x_n')]}\}.$$

Set $\varphi(t) = E(e^{itx})$, $\psi(t) = E(e^{itx'})$. The relation above becomes

$$\varphi(t_1)\varphi(t_2)\varphi(-t_1-t_2) = \psi(t_1)\psi(t_2)\psi(-t_1-t_2). \tag{1}$$

Consider equation (1) for values of t_1, t_2 in the neighborhood of $t = 0$. $\varphi(0) = \psi(0) = 1$, hence there is an interval $|t| < \delta$, in which $\varphi(t)$ and $\psi(t)$ do not vanish. It is easily shown that $\varphi(t)$ and $\psi(t)$ are each continuous, since e^{itx}, in the neighborhood of $t = 0$, is continuous uniformly for any bounded interval of x, and since A may be chosen so that $1 - F(A)$ and $F(-A)$ are both as small as desired. In the interval $|t| < \delta$ the function $f(t) = \varphi(t)/\psi(t)$ is continuous. Also, $\varphi(-t) = \overline{\varphi(t)}$ and $\psi(-t) = \overline{\psi(t)}$. Setting $t_2 = 0$ in (1) we obtain $\varphi(t)\varphi(-t) = \psi(t)\psi(-t)$, hence $|\varphi(t)| = |\psi(t)|$, that is, $|f(t)| = 1$. $f(t)$ takes values on the unit circle of the complex plane, and $f(0) = 1$, hence there is an interval $|t| < \delta'$ such that $z = f(t)$ lies on an arc γ, of length less than 2π, containing the point $z = 1$. Now consider the functional equation (1) for $|t_1| < \frac{1}{2}\delta'$, $|t_2| < \frac{1}{2}\delta'$. (1) becomes

$$f(t_1)f(t_2)f(-t_1-t_2) = 1.$$

The interval $|t| < \delta'$ was so chosen that for $|t_1| < \frac{1}{2}\delta'$, $|t_2| < \frac{1}{2}\delta'$, it is possible to define a single-valued branch of the argument of $f(t_1)$, $f(t_2)$, and $f(t_1 + t_2)$. Letting $t_2 = 0$ we have $f(t)f(-t) = 1$, hence, replacing $f(-t_1-t_2)$ by $1/f(t_1 + t_2)$ in the last equation, we have

$$f(t_1)f(t_2) = f(t_1 + t_2).$$

Arg $f(t_1)$, arg $f(t_2)$, and arg $f(t_1 + t_2)$ are uniquely determined, except for some *fixed* multiple of 2π. If we choose the principal value of the argument, i.e., so

[2] The set S is ϵ-connected if any two points p, q, in S can be connected by an ϵ-chain, i.e., there exists a set $p_0 = p, p_1, \cdots, p_{n-1}, p_n = q$, such that $|p_i - p_{i-1}| < \epsilon$, $i = 1, 2, \cdots, n$.

that $0 \leq \arg f(t) < 2\pi$, we must have

$$\arg f(t_1) + \arg f(t_2) = \arg f(t_1 + t_2)$$

for $|t_1| < \frac{1}{2}\delta'$, $|t_2| < \frac{1}{2}\delta'$. Since $\arg f(t)$ is continuous, any solution of this well known functional equation must be of the form $\arg f(t) = at$. $|f(t)| = 1$, therefore there exists a constant a such that $f(t) = e^{iat}$, for $|t| < \frac{1}{2}\delta'$, that is, $\varphi(t) = e^{iat}\psi(t)$, for $|t| < \frac{1}{2}\delta'$. By use of (1) this may be extended to hold for all t such that $|t| < \epsilon$, where ϵ is the minimum modulus of all t such that $\varphi(t) = 0$. (1) may now be used to extend the relation for all t such that $\varphi(t) \neq 0$ by choosing an ϵ-chain connecting the origin to the point t. We know already that $\varphi(t) = e^{iat}\psi(t)$ if $\varphi(t) = 0$, hence it holds for all t. This relation says that $E(e^{itx}) = E(e^{it(x'+a)})$, hence $x' + a$ and x have the same distribution, thus completing the demonstration of the theorem.

It should be remarked that the set $(x_1 - x_n, \cdots, x_{n-1} - x_n)$ may be replaced in Theorem Ia by any equivalent set, for example, $(x_1 - \bar{x}, \cdots, x_{n-1} - \bar{x})$.

The next result is of the same nature as Theorem Ia except for the replacement of the location parameter by a scale (positive or negative) parameter.

THEOREM IB: *Let x have distribution $F(x)$, such that the zeros of $\int_{-\infty}^{\infty} e^{it(\log|x|)}\, dF(x)$ are nowhere dense, and let x' have distribution $G(x')$. Let $x_1, \cdots, x_n$ and $x_1', \cdots, x_n'$ be samples from the distributions of x and x', with $n \geq 3$, then the set $w_\alpha = x_\alpha/x_n$, $\alpha = 1, \cdots, n-1$, have the same distribution as the set $w_\alpha' = x_\alpha'/x_n'$ if and only if there exists a constant c such that cx' and x have the same distribution.*

PROOF: The sufficiency of the condition is evident. Suppose, then, as before, that w_1, w_2 have the same joint distribution as w_1', w_2'. Log $|w_1|$ and log $|w_2|$ have the same joint distribution as log $|w_1'|$ and log $|w_2'|$, hence by application of Theorem Ia to log $|x|$ and log $|x'|$ it follows (since the complement of a nowhere dense set is ϵ-connected for every ϵ) that there exists a constant a such that

$$\int_{-\infty}^{\infty} e^{it\log|x|}\, dF(x) = \int_{-\infty}^{\infty} e^{it[\log|x'|-a]}\, dG(x).$$

Let $y = e^{-a}x'$, then $|x|$ and $|y|$ have the same distribution, and

$$\int e^{it\log|x|}\, dF(x) = \int e^{it\log|y|}\, dH(y), \tag{2}$$

where y has distribution $H(y)$. We now have to show that either y or $-y$ has the distribution of x, that is, it must be shown that either $H(y) = F(y)$, or $H(y) = 1 - F(-y)$.

By the first part of the theorem the functions $u_1 = y_1/y_3$ and $u_2 = y_2/y_3$ have the same joint distribution as w_1, w_2. It is clear that the mean value of any function of u_1 and u_2 is the same as the mean value of the corresponding function of w_1 and w_2. Hence

$$\iiint_{-\infty}^{\infty} e^{i[t_1 \log|w_1| + t_2 \log|w_2|]} \operatorname{sgn} w_1 \operatorname{sgn} w_2 \, dF(x_1)\, dF(x_2)\, dF(x_3)$$

$$= \iiint_{-\infty}^{\infty} e^{i[t_1 \log|u_1| + t_2 \log|u_2|]} \operatorname{sgn} u_1 \operatorname{sgn} u_2 \, dH(y_1)\, dH(y_2)\, dH(y_3),$$

where sgn $x = 1$, for $x \geq 0$, sgn $x = -1$ for $x < 0$.

$$(\operatorname{sgn} w_1)(\operatorname{sgn} w_2) = (\operatorname{sgn} x_1)(\operatorname{sgn} x_2),$$

so that the last equation becomes

$$\iiint_{-\infty}^{\infty} e^{i[t_1 (\log|x_1| - \log|x_3|) + t_2 (\log|x_2| - \log|x_3|)]} \operatorname{sgn} x_1 \operatorname{sgn} x_2 \, dF(x_1)\, dF(x_2)\, dF(x_3)$$

$$(3) \qquad = \iiint_{-\infty}^{\infty} e^{i[t_1 (\log|y_1| - \log|y_3|) + t_2 (\log|y_2| - \log|y_3|)]} \operatorname{sgn} y_1 \times \operatorname{sgn} y_2 \, dH(y_1)\, dH(y_2)\, dH(y_3).$$

Set

$$\psi_1(t) = \int e^{it \log|x|} dF(x); \qquad \varphi_1(t) = \int e^{it \log|y|} dH(y)$$

$$\psi_2(t) = \int e^{it \log|x|} \operatorname{sgn} x \, dF(x); \qquad \varphi_2(t) = \int e^{it \log|y|} \operatorname{sgn} y \, dH(y).$$

From (3) we have $\psi_2(t_1)\psi_2(t_2)\psi_1(-t_1 - t_2) = \varphi_2(t_1)\varphi_2(t_2)\varphi_1(-t_1 - t_2)$ for all t_1, t_2, and from (2) we have $\psi_1(t) = \varphi_1(t)$ for all t, hence, if $\psi_1(-t_1 - t_2) \neq 0$, $\psi_2(t_1)\psi_2(t_2) = \varphi_2(t_1)\varphi_2(t_2)$. By hypothesis the zeros of $\psi_1(t)$ are nowhere dense, hence if $\psi_1(-t_1 - t_2) = 0$ there is a sequence $t^{(n)}$, such that $t^{(n)} \to -t_1 - t_2$ and $\psi_1(t^{(n)}) \neq 0$. Now take an arbitrary sequence $t_1^{(n)}$ such that $t_1^{(n)} \to t_1$, then $t_2^{(n)} = -t^{(n)} - t_1^{(n)}$ must tend to t_2. For each n we have $\psi_2(t_1^{(n)})\psi_2(t_2^{(n)}) = \varphi_2(t_1^{(n)})\varphi_2(t_2^{(n)})$. All the functions appearing are continuous, thus we see that $\psi_2(t_1)\psi_2(t_2) = \varphi_2(t_1)\varphi_2(t_2)$ for all t_1, t_2. From this it follows directly that either $\psi_2(t) = \varphi_2(t)$ for all t or $\psi_2(t) = -\varphi_2(t)$ for all t. We have[3]

$$\psi_1(t) = \int_0^{\infty} e^{it \log x} dF(x) + \int_{-\infty}^{0} e^{it \log(-x)} dF(x)$$

$$\psi_2(t) = \int_0^{\infty} e^{it \log x} dF(x) - \int_{-\infty}^{0} e^{it \log(-x)} dF(x)$$

[3] The assumption has been made implicitly that $F(x)$ and $G(x)$ are continuous at $x = 0$, otherwise the distribution of x_i/x_n is not properly defined, and the functions $\varphi_i(t)$ and $\psi_i(t)$ are then not defined. Similar assumptions will be made whenever necessary in later theorems.

$$\varphi_1(t) = \int_0^\infty e^{it\log x}\, dH(x) + \int_{-\infty}^0 e^{it\log(-x)}\, dH(x)$$

and
$$\varphi_2(t) = \int_0^\infty e^{it\log x}\, dH(x) - \int_{-\infty}^0 e^{it\log(-x)}\, dH(x).$$

Combining these expressions with the relations obtained above leads, by Fourier inversion, to the result that either $F(x) \equiv H(x)$ or $H(x) \equiv 1 - F(-x)$. We have shown that either y or $-y$ has the same distribution as x, that is, either $e^{-a}x'$ or $-e^{-a}x'$ has the same distribution as x.

Theorem Ib states essentially that the joint distribution of the set x_α/x_n, $\alpha = 1, \cdots, n-1$, determines the distribution of x except for a scale parameter and possibly a reflection. In the event that x has an asymmetrical distribution, and if it is desired to rule out negative changes of scale, a variation of this procedure is necessary. The next result is appropriate for this situation.

Theorem Ic: *Let x have distribution $F(x)$ such that the zeros of $\int e^{it\log|x|}\, dF(x)$ are nowhere dense, and let x' have distribution $G(x')$. Let $x_1, \cdots, x_n$ and $x_1', \cdots, x_n'$ be samples from the distributions of x and x', with $n \geq 3$. Express $x_1, \cdots, x_n$ and $x_1', \cdots, x_n'$ in spherical coordinates*

$$\begin{aligned} x_1 &= r\cos\theta_1, & x_1' &= r'\cos\theta_1' \\ x_2 &= r\sin\theta_1\cos\theta_2, & x_2' &= r'\sin\theta_1'\cos\theta_2' \\ &\vdots & &\vdots \\ x_n &= r\sin\theta_1\sin\theta_2\cdots\sin\theta_{n-1}, & x_n' &= r'\sin\theta_1'\sin\theta_2'\cdots\sin\theta_{n-1}'. \end{aligned}$$

Then $\theta_1, \cdots, \theta_{n-1}$ have the same joint distribution as $\theta_1', \cdots, \theta_{n-1}'$ if and only if there exists a positive constant k such that kx' and x have the same distribution.

Proof: Sufficiency of the condition is an immediate consequence of the fact that $\theta_1, \cdots, \theta_{n-1}$ are invariant under the transformation $x = kx'$, with $k > 0$. If $\theta_1, \cdots, \theta_{n-1}$ have the same joint distribution as $\theta_1', \cdots, \theta_{n-1}'$ then the set $\{x_\alpha/x_n\}$ have the same joint distribution as the set $\{x_\alpha'/x_n'\}$, hence, by Theorem Ib, there exists a constant c such that cx' has the same distribution as x. To establish necessity of the condition we must show that $|c|x'$ has the same distribution as x.

Set $y = |c|x'$, and let $y_1, \cdots, y_n$ be expressed in spherical coordinates; $y_1, \cdots, y_n$ have the same angular coordinates $\theta_1', \cdots, \theta_{n-1}'$. This implies that x_1/r and x_2/r have the same joint distribution as y_1/R and y_2/R, where $R = \sqrt{y_1^2 + \cdots + y_n^2}$; $\dfrac{x_1}{r}\Big/\left|\dfrac{x_2}{r}\right| = x_1/|x_2|$, therefore $x_1/|x_2|$ has the same distribution as $y_1/|y_2|$, so that

$$\iint_{-\infty}^{\infty} e^{it\log\left|\frac{x_1}{x_2}\right|}\operatorname{sgn}\left(\frac{x_1}{|x_2|}\right) dF(x_1)\, dF(x_2) = \iint_{-\infty}^{\infty} e^{it\log\left|\frac{y_1}{y_2}\right|}\operatorname{sgn}\left(\frac{y_1}{|y_2|}\right) dH(y_1)\, dH(y_2)$$

if y has distribution $H(y)$. $\operatorname{Sgn}\left(\frac{x_1}{|x_2|}\right) = \operatorname{sgn} x_1$, so that the last equation yields

$$\int_{-\infty}^{\infty} e^{it \log |x|} \operatorname{sgn} x \, dF(x) \cdot \int_{-\infty}^{\infty} e^{-it \log |x|} \, dF(x) = \int_{-\infty}^{\infty} e^{it \log |x|} \operatorname{sgn} x \, dH(x) \cdot \int_{-\infty}^{\infty} e^{-it \log |x|} \, dH(x).$$

We know already that $|x|$ and $|y|$ have the same distribution, so that

$$(4) \qquad \int_{-\infty}^{\infty} e^{it \log |x|} \, dF(x) = \int_{-\infty}^{\infty} e^{it \log |x|} \, dH(x),$$

thus

$$(5) \qquad \int_{-\infty}^{\infty} e^{it \log |x|} \operatorname{sgn} x \, dF(x) = \int_{-\infty}^{\infty} e^{it \log |x|} \operatorname{sgn} x \, dH(x),$$

except possibly for zeros of $\int_{-\infty}^{\infty} e^{-it \log |x|} \, dF(x)$. By hypothesis the exceptional points are nowhere dense, so that, by continuity, (5) holds for all t. (4) and (5) together imply, as in the proof of Theorem Ib, that $F(x) \equiv H(x)$, i.e., x and $|c| x'$ have the same distribution.

The next three results are generalizations of Theorems Ia, b, c, to analogous multivariate situations. The first of these is a direct generalization of Theorem Ia.

THEOREM IIA: *Let $x_1, \cdots, x_k$ have joint distribution $F(x_1, \cdots, x_k)$ such that the complement of the set S of zeros of $\int e^{i\Sigma t_r x_r} dF(x_1, \cdots, x_k)$ is ϵ-connected, where ϵ is the* g.l.b. *of $|t|$ for (t) in S, and let $y_1, \cdots, y_k$ have joint distribution $G(y_1, \cdots, y_k)$. Let $(x_1^\alpha, \cdots, x_k^\alpha)$ and $(y_1^\alpha, \cdots, y_k^\alpha)$, $\alpha = 1, \cdots, n$, be samples from these distributions, with $n \geq 3$. Then $w_{i\beta} = x_i^\beta - x_i^n$, $i = 1, \cdots, k$, $\beta = 1, \cdots, n - 1$, have the same joint distribution as the corresponding set $v_{i\beta} = y_i^\beta - y_i^n$ if and only if there exist constants $a_1, \cdots, a_k$ such that $y_1 + a_1, \cdots, y_k + a_k$ have the same joint distribution as $x_1, \cdots, x_k$.*

PROOF: Set

$$\varphi(t_1, \cdots, t_k) = \int e^{i \sum_1^k t_r x_r} dF(x_1, \cdots, x_k),$$

$$\psi(t_1, \cdots, t_k) = \int e^{i \sum_1^k t_r y_r} dG(y_1, \cdots, y_k).$$

If $w_{i\beta}$, $i = 1, \cdots, k$, $\beta = 1, 2$, have the same joint distribution as $v_{i\beta}$, then, as in the proof of Theorem Ia, we have

$$(6) \qquad \begin{aligned} &\varphi(t_{11}, \cdots, t_{k1})\varphi(t_{12}, \cdots, t_{k2})\varphi(-t_{11} - t_{12}, \cdots, -t_{k1} - t_{k2}) \\ &\quad = \psi(t_{11}, \cdots, t_{k1})\psi(t_{12}, \cdots, t_{k2})\psi(-t_{11} - t_{12}, \cdots, -t_{k1} - t_{k2}). \end{aligned}$$

Again, as before, $|\varphi| = |\psi|$; $\varphi(t_1, \cdots, t_k)$ and $\psi(t_1, \cdots, t_k)$ are continuous; $\varphi(0, 0, \cdots, 0) = \psi(0, 0, \cdots, 0) = 1$. There will exist a neighborhood N of $(0, 0, \cdots, 0)$ such that for $(t_1, \cdots, t_k) \in N$ the function $f(t_1, \cdots, t_k) = \dfrac{\varphi(t_1, \cdots, t_k)}{\psi(t_1, \cdots, t_k)}$ is defined and continuous. Then there will exist a neighborhood $N' \subset N$ such that in N' there exists a uniquely determined branch of $\arg f(t_1, \cdots, t_k)$, continuous in N', and such that if $(t_1, \cdots, t_k) \in N'$ and $(u_1, \cdots, u_k) \in N'$ then $\arg f(t_1 + u_1, \cdots, t_k + u_k)$ is also uniquely determined and continuous. For $(t) \in N'$ and $(u) \in N'$, $\arg f$ satisfies the relation

$$\arg f(t_1, \cdots, t_k) + \arg f(u_1, \cdots, u_k) = \arg f(t_1 + u_1, \cdots, t_k + u_k).$$

It is easily shown that any continuous function satisfying the equation above must be of the form $\Sigma a_r t_r$, therefore

$$\varphi(t_1, \cdots t_k) = e^{i \sum_1^k a_r t_r} \psi(t_1, \cdots, t_k); \qquad (t) \in N'. \tag{7}$$

Just as in the proof of Ia the relation (7) may be extended, by use of (6), to hold for all t. This implies, finally, that the set $\{y_i + a_i\}$ have the same joint distribution as the set $\{x_i\}$.

Theorem IIb is a generalization of Theorem Ib to multivariate distributions.

THEOREM IIB: *Let $x_1, \cdots, x_k$ have distribution $F(x_1, \cdots, x_k)$ such that the zeros of $\int e^{i\Sigma t_r \log |x_r|} dF(x_1, \cdots, x_k)$ are nowhere dense, and let $y_1, \cdots, y_k$ have distribution $G(y_1, \cdots, y_k)$. Let $(x_1^\alpha, \cdots, x_k^\alpha)$ and $(y_1^\alpha, \cdots, y_k^\alpha)$, $\alpha = 1, \cdots, n$, be samples, with $n \geq 3$. Then the set $w_{i\beta} = x_i^\beta / x_i^n$, $i = 1, \cdots, k$, $\beta = 1, \cdots, n - 1$, have the same joint distribution as the corresponding set $v_{i\beta} = y_i^\beta / y_i^n$ if and only if there exist constants $c_1, \cdots, c_k$ such that the set $c_i y_i$ have the same distribution as the x_i.*

PROOF: The demonstration is parallel to that of Theorem Ib. By Theorem IIa there exist $a_1, \cdots, a_k$ such that

$$E(e^{i\Sigma t_r \log |x_r|}) = E(e^{i\Sigma t_r (\log |y_r| + a_r)}).$$

Set $z_r = e^{a_r} y_r$, then

$$\int e^{i\Sigma t_r \log |x_r|} dF(x_1, \cdots, x_k) = \int e^{i\Sigma t_r \log |z_r|} dH(z_1, \cdots, z_k), \tag{8}$$

where $(z_1, \cdots, z_k)$ have distribution function $H(z_1, \cdots, z_k)$.

We shall continue the proof from here under the assumption that $k = 2$. It will be evident how the proof goes for any k. We have, since z_r^β / z_i^3 have the same joint distribution as x_r^β / x_i^3,

$$\begin{aligned} &\iiint_{-\infty}^{\infty} e^{i\Sigma t_{r\beta}(\log |x_r^\beta| - \log |x_r^3|)} \operatorname{sgn}\left(\frac{x_1^1}{x_1^3}\right) \operatorname{sgn}\left(\frac{x_1^2}{x_1^3}\right) dF(x_1^1, x_2^1)\, dF(x_1^2, x_2^2)\, dF(x_1^3, x_2^3) \\ &= \iiint_{-\infty}^{\infty} e^{i\Sigma t_{r\beta}(\log |x_r^\beta| - \log |x_r^3|)} \operatorname{sgn}\left(\frac{x_1^1}{x_1^3}\right) \operatorname{sgn}\left(\frac{x_1^2}{x_1^3}\right) dH(x_1^1, x_2^1)\, dH(x_1^2, x_2^2)\, dH(x_1^3, x_2^3). \end{aligned} \tag{9}$$

Both members of (9) are evaluated as products, just as was done in previous proofs, and from the result, combined with (8), we conclude, as in Theorem Ib, that

$$\iint_{-\infty}^{\infty} e^{i\Sigma t_r \log|x_r|} \operatorname{sgn} x_1 \, dF(x_1, x_2) = s_1 \iint_{-\infty}^{\infty} e^{i\Sigma t_r \log|x_r|} \operatorname{sgn} x_1 \, dH(x_1, x_2),$$

where $s_1 = \pm 1$, for all (t_1, t_2). Similarly

$$\iint_{-\infty}^{\infty} e^{i\Sigma t_r \log|x_r|} \operatorname{sgn} x_2 \, dF(x_1, x_2) = s_2 \iint_{-\infty}^{\infty} e^{i\Sigma t_r \log|x_r|} \operatorname{sgn} x_2 \, dH(x_1, x_2)$$

and

$$\iint_{-\infty}^{\infty} e^{i\Sigma t_r \log|x_r|} \operatorname{sgn} x_1 \operatorname{sgn} x_2 \, dF(x_1, x_2) = s_3 \iint_{-\infty}^{\infty} e^{i\Sigma t_r \log|x_r|} \operatorname{sgn} x_1 \operatorname{sgn} x_2 \, dH(x_1, x_2),$$

with $s_2 = \pm 1$, $s_3 = \pm 1$.

Set
$$\varphi_1(t_1, t_2) = \iint_{-\infty}^{\infty} e^{i\Sigma t_r \log|x_r|} \operatorname{sgn} x_1 \, dF(x_1, x_2)$$

$$\varphi_2(t_1, t_2) = \iint_{-\infty}^{\infty} e^{i\Sigma t_r \log|x_r|} \operatorname{sgn} x_2 \, dF(x_1, x_2)$$

$$\varphi_{12}(t_1, t_2) = \iint_{-\infty}^{\infty} e^{i\Sigma t_r \log|x_r|} \operatorname{sgn} x_1 \operatorname{sgn} x_2 \, dF(x_1, x_2)$$

and let $\psi_1(t_1, t_2)$, $\psi_2(t_1, t_2)$, and $\psi_{12}(t_1, t_2)$ denote the corresponding transforms of $H(x_1, x_2)$. We have

$$(10) \qquad \begin{cases} \varphi_1(t_1, t_2) = s_1 \psi_1(t_1, t_2) \\ \varphi_2(t_1, t_2) = s_2 \psi_2(t_1, t_2) \\ \varphi_{12}(t_1, t_2) = s_3 \psi_{12}(t_1, t_2) \end{cases}$$

with $s_1 = \pm 1$, $s_2 = \pm 1$, and $s_3 = \pm 1$.

Now, as in (9), by considering $E\left[e^{i\Sigma t_{r\beta}(\log|x_r^\beta| - \log|x_r^3|)} \operatorname{sgn}\left(\frac{x_1^1}{x_1^3}\right) \operatorname{sgn}\left(\frac{x_2^2}{x_2^3}\right)\right]$ we obtain the relation

$$\varphi_1(t_{11}, t_{21})\varphi_2(t_{12}, t_{22})\varphi_{12}(-t_{11} - t_{12}, -t_{21} - t_{22})$$
$$= \psi_1(t_{11}, t_{21})\psi_2(t_{12}, t_{22})\psi_{12}(-t_{11} - t_{12}, -t_{21} - t_{22}),$$

showing that s_1, s_2, s_3, may be chosen so that $s_1 s_2 s_3 = 1$, that is, $s_1 s_2 = s_3$.

Consider now the variates $z_r' = s_r z_r$, $r = 1, 2$. Let $K(z_1', z_2')$ be the distribution function of z_1', z_2'. If we let $\theta_1(t_1, t_2)$, $\theta_2(t_1, t_2)$, and $\theta_{12}(t_1, t_2)$ be the transforms of K which correspond to $\varphi_1(t_1, t_2)$, $\varphi_2(t_1, t_2)$, and $\varphi_{12}(t_1, t_2)$ respectively, it is evident that

$$\begin{aligned} \varphi_1(t_1, t_2) &= \theta_1(t_1, t_2) \\ \varphi_2(t_1, t_2) &= \theta_2(t_1, t_2) \\ \varphi_{12}(t_1, t_2) &= \theta_{12}(t_1, t_2). \end{aligned} \tag{11}$$

Moreover, from (8),

$$\iint_{-\infty}^{\infty} e^{i\Sigma t_r \log |x_r|} \, dF(x_1, x_2) = \iint_{-\infty}^{\infty} e^{i\Sigma t_r \log |x_r|} \, dK(x_1, x_2).$$

The last relation, together with the equations (11) imply that $F(x)$ and $K(x)$ coincide in each quadrant, thus $F(x_1, x_2) \equiv K(x_1, x_2)$ for all x_1, x_2.

The final result is that z_1', z_2' have the same distribution as x_1, x_2, i.e., $s_1 e^{a_1} y_1$ and $s_2 e^{a_2} y_2$ have the same joint distribution as x_1 and x_2.

The next result bears the same relation to Theorem IIb that Theorem Ic bears to Theorem Ib, that is, only *positive* scale changes are to be permitted.

THEOREM IIc: *Let $x_1, \cdots, x_k$ have distribution $F(x_1, \cdots, x_k)$ such that the zeros of $\int e^{i\Sigma t_r \log |x_r|} \, dF(x_1, \cdots, x_k)$ are nowhere dense, and let $y_1, \cdots, y_k$ have distribution $G(y_1, \cdots, y_k)$. Let $(x_1^\alpha, \cdots, x_k^\alpha)$ and $(y_1^\alpha, \cdots, y_k^\alpha)$, $\alpha = 1, 2, \cdots, n$, be samples with $n \geq 3$. Express $x_1^\alpha, \cdots, x_k^\alpha$ and $y_1^\alpha, \cdots, y_k^\alpha$ in spherical coordinates*

$$\begin{aligned} x_i^1 &= r_i \cos \theta_i^1, & y_i^1 &= R_i \cos \varphi_i^1, \\ x_i^2 &= r_i \sin \theta_i^1 \cos \theta_i^2, & y_i^2 &= R_i \sin \varphi_i^1 \cos \varphi_i^2, \\ &\vdots & &\vdots \\ x_i^n &= r_i \sin \theta_i^1 \cdots \sin \theta_i^{n-1}; & y_i^n &= R_i \sin \varphi_i^1 \cdots \sin \varphi_i^{n-1}. \end{aligned}$$

Then $\{\theta_i^\beta\}$, $i = 1, \cdots, k$, $\beta = 1, \cdots, n-1$, have the same joint distribution as $\{\varphi_i^\beta\}$ if and only if there exist constants $k_i > 0$, $i = 1, \cdots, k$, such that the set $k_i y_i$ have the same joint distribution as the set x_i.

PROOF: If $\{\theta_i^\beta\}$ have the same distribution as $\{\varphi_i^\beta\}$ then it follows that $\left\{\frac{x_i^\beta}{x_i^n}\right\}$ have the same distribution as $\left\{\frac{y_i^\beta}{y_i^n}\right\}$, hence by Theorem IIb there exist constants c_i such that $\{c_i y_i\}$ have the same distribution as $\{x_i\}$. Set $z_i = |c_i| y_i$; we wish to show that $\{z_i\}$ have the same distribution as $\{x_i\}$. By equation (8) in Theorem IIb it is known that $\{|z_i|\}$ have the same distribution as $\{|x_i|\}$, moreover, if we express z_i^α in spherical coordinates, the angular coordinates are the same as those of y_i^α, therefore $\left\{\frac{x_i^1}{|x_i^2|}\right\}$ have the same distribution as $\left\{\frac{z_i^1}{|z_i^2|}\right\}$, since these functions are obtainable in terms of the angular coordinates.

As before, we shall continue the proof from here under the assumption that $k = 2$. The procedure is a generalization of the procedure in the proof of Theorem Ic. $\operatorname{sgn} x_i^1 = \operatorname{sgn}\left\{\frac{x_i^1}{|x_i^2|}\right\}$, and similarly for y, therefore

$$(12)\quad \begin{aligned} &\iint e^{i\sum_{r=1}^{2} t_r(\log|x_r^1|-\log|x_r^2|)} \operatorname{sgn} x_i^1\, dF(x_1^1, x_2^1)\, dF(x_1^2, x_2^2) \\ &= \iint e^{i\sum_{r=1}^{2} t_r(\log|x_r^1|-\log|x_r^2|)} \operatorname{sgn} x_i^1\, dH(x_1^1, x_2^1)\, dH(x_1^2, x_2^2), \qquad i = 1, 2, \end{aligned}$$

where it is assumed that z_1, z_2 have distribution $H(z_1, z_2)$. As before, set

$$\varphi(t_1, t_2) = \int e^{i\Sigma t_r \log|x_r|}\, dF(x_1, x_2),$$

$$\varphi_i(t_1, t_2) = \int e^{i\Sigma t_r \log|x_r|} \operatorname{sgn} x_i\, dF(x_1, x_2), \qquad i = 1, 2,$$

$$\varphi_{12}(t_1, t_2) = \int e^{i\Sigma t_r \log|x_r|} \operatorname{sgn} x_1 \operatorname{sgn} x_2\, dF(x_1, x_2),$$

and denote the corresponding transforms of $H(x_1, x_2)$ by $\theta(t_1, t_2)$, $\theta_1(t_1, t_2)$, $\theta_2(t_1, t_2)$, and $\theta_{12}(t_1, t_2)$. It has been remarked already that $\{|z_i|\}$ have the same distribution as $\{|x_i|\}$, therefore $\theta(t_1, t_2) = \varphi(t_1, t_2)$. Equation (12) yields the relation $\varphi_i(t_1, t_2)\varphi(-t_1, -t_2) = \theta_i(t_1, t_2)\theta(-t_1, -t_2)$, $i = 1, 2$; the zeros of $\varphi(t_1, t_2)$ are nowhere dense, so that it can be concluded that $\varphi_i(t_1, t_2) = \theta_i(t_1, t_2)$, $i = 1, 2$. Now, from an equation similar to (12) we obtain $\varphi_{12}(t_1, t_2) = \theta_{12}(t_1, t_2)$. As in Theorem IIb, the four relations above together imply that $F(x_1, x_2) \equiv H(x_1, x_2)$, in other words, $\{|c_i| y_i\}$ have the same distribution as $\{x_i\}$.

We are now in a position to combine some of the preceding theorems so as to obtain analogous results for scale and location parameters together.

THEOREM IIIA: *Let x have distribution $F(x)$ such that the zeros of $\int e^{itx}\, dF(x)$ satisfy the condition of Theorem* Ia, *and the zeros of*

$$\iiint e^{it_1 \log|x_1-x_3|+it_2 \log|x_2-x_3|}\, dF(x_1)\, dF(x_2)\, dF(x_3)$$

are nowhere dense, and let y have distribution $G(y)$. Let $x_1, \cdots, x_n$ and $y_1, \cdots, y_n$ be samples, with $n \geq 9$. Then $w_\alpha = \frac{x_\alpha - x_n}{x_{n-1} - x_n}$, $\alpha = 1, \cdots, n-2$, have the same joint distribution as the corresponding set $w'_\alpha = \frac{y_\alpha - y_n}{y_{n-1} - y_n}$ if and only if there exist constants a, c, such that $c(y - a)$ and x have the same distribution.

PROOF: Sufficiency of the condition is an immediate consequence of the fact that w'_α is invariant under transformations of the form $y' = c(y - a)$. Assume then that $\{w_\alpha\}$ and $\{w'_\alpha\}$ have the same joint distribution. By elementary transformations it is evident that the functions $\frac{x_1 - x_3}{x_7 - x_9}, \frac{x_4 - x_6}{x_7 - x_9}, \frac{x_2 - x_3}{x_8 - x_9}, \frac{x_5 - x_6}{x_8 - x_9}$, have the same joint distribution as the corresponding functions of the y's, if $n \geq 9$. Since $x_1, \cdots, x_n$ form a sample it follows that the pairs $\{x_1 - x_3, x_2 - x_3\}$, $\{x_4 - x_6, x_5 - x_6\}$, $\{x_7 - x_9, x_8 - x_9\}$, have the same joint distributions and are pairwise independent, and similarly for the corresponding functions of the y's. Theorem IIb assures the existence of constants c_1, c_2, such that $c_1(y_1 - y_3)$, $c_2(y_2 - y_3)$ have the same joint distribution as $(x_1 - x_3)$, $(x_2 - x_3)$. Considering separately the marginal distributions it is seen that $c_1(y_1 - y_3)$ has the same distribution as $c_2(y_2 - y_3)$. $y_1 - y_3$ and $y_2 - y_3$ have the same distribution, therefore either $c_2 = c_1$, or $c_2 = -c_1$. Set $u_\alpha = x_\alpha - x_3$, $v_\alpha = c_1(y_\alpha - y_3)$, $\alpha = 1, 2$. We have, for the distributions of (u_1, u_2) and (v_1, v_2), relations corresponding to (10) in Theorem IIb, with the additional condition that $s_1 = s_2$, because of the symmetry in the variables. This implies that either (v_1, v_2) or $(-v_1, -v_2)$ have the same joint distribution as (u_1, u_2), that is, there exists c such that $c(y_1 - y_3)$ and $c(y_2 - y_3)$ have the same joint distribution as $x_1 - x_3$ and $x_2 - x_3$. Application of Theorem Ia now completes the proof.

Just as before, there is an analogous situation when we consider angular coordinates instead of quotients. The proof is immediate; the angular coordinates determine the angular coordinates of $\{x_1 - x_3, x_2 - x_3\}$, $\{x_4 - x_6, x_5 - x_6\}$, and $\{x_7 - x_9, x_8 - x_9\}$, arranged as a sample. Then the constants c_1, c_2 in the proof of Theorem IIIa are both positive; it follows that $c_1 = c_2$. Application of Theorem Ia gives

THEOREM IIIB: *Let $x_1, \cdots, x_n$ and $y_1, \cdots, y_n$ satisfy the hypotheses of Theorem* IIIa. *Set*

$$x_1 - x_n = r \cos \theta_1, \qquad y_1 - y_n = r' \cos \theta'_1,$$

$$x_2 - x_n = r \sin \theta_1 \cos \theta_2, \qquad y_2 - y_n = r' \sin \theta'_1 \cos \theta'_2,$$

$$x_{n-1} - x_n = r \sin \theta_1 \cdots \sin \theta_{n-2}; \qquad y_{n-1} - y_n = r' \sin \theta'_1 \cdots \sin \theta'_{n-2}.$$

Then $\theta_1, \cdots \theta_{n-2}$ have the same joint distribution as $\theta'_1, \cdots, \theta'_{n-2}$ if and only if there exist constants a and $c > 0$ such that $c(y - a)$ has the same distribution as x.

Theorem IVa is a generalization of Theorem Ia to cover arbitrary linear combinations of some subset of the sample.

THEOREM IVA: *Suppose x has distribution $F(x)$ such that $\int e^{itx} dF(x)$ does not vanish, and let y have distribution $G(y)$. Consider the functions* $w_\alpha = x_\alpha - \sum_{\beta=1}^{n-m} l_{\alpha\beta} x_{m+\beta}$, $w'_\alpha = y_\alpha - \sum_{\beta=1}^{n-m} l_{\alpha\beta} y_{m+\beta}$, $\alpha = 1, 2, \cdots, m$, $\beta = 1, 2, \cdots$;

$n - m$, and suppose that $m > n - m$. Then, if $\{w_\alpha\}$ have the same joint distribution as $\{w'_\alpha\}$ and if $\sum_{\beta=1}^{n-m} l_{\alpha\beta} \neq 1$ for some α, it follows that $F(y) \equiv G(y)$; if $\sum_\beta l_{\alpha\beta} = 1$ for all α there exists a constant a such that $F(y - a) \equiv G(y)$.

PROOF: Denote the characteristic functions of x and y by $\varphi(t)$ and $\psi(t)$ respectively. By expressing the fact that $\{w_\alpha\}$ and $\{w'_\alpha\}$, $\alpha = 1, 2, \cdots, n - m + 1$, have the same characteristic function we obtain the functional equation

$$\prod_{\alpha=1}^{n-m+1} \varphi(t_\alpha) \prod_{\beta=1}^{n-m} \varphi\left(-\sum_{\alpha=1}^{n-m+1} l_{\alpha\beta} t_\alpha\right) = \prod_{\alpha=1}^{n-m+1} \psi(t_\alpha) \prod_{\beta=1}^{n-m} \psi\left(-\sum_{\alpha=1}^{n-m+1} l_{\alpha\beta} t_\alpha\right).$$

By hypothesis $\varphi(t)$ does not vanish, therefore $\psi(t)$ has no zeros, because of the relation above. $\varphi(t)$ and $\psi(t)$ are continuous, thus the function $f(t) = \log \varphi(t) - \log \psi(t)$ can be uniquely defined in a continuous manner for all t. The equation above becomes

$$\sum_{\alpha=1}^{n-m+1} f(t_\alpha) + \sum_{\beta=1}^{n-m} f\left(-\sum_{\alpha=1}^{n-m+1} l_{\alpha\beta} t_\alpha\right) = 0. \tag{13}$$

The constants $l_{\alpha\beta}$ are necessarily linearly dependent, so that, for some α, $l_{\alpha\beta}$ can be expressed as a linear combination of the others; suppose then that

$$l_{n-m+1,\beta} = \sum_{\alpha=1}^{n-m} e_\alpha l_{\alpha\beta}.$$

Putting these values in (13) we have

$$\sum_{\alpha=1}^{n-m+1} f(t_\alpha) + \sum_{\beta=1}^{n-m} f\left(-\sum_{\alpha=1}^{n-m} l_{\alpha\beta}(t_\alpha + t_{n-m+1} c_\alpha)\right) = 0. \tag{14}$$

It can be assumed that $\Sigma e_\alpha^2 \neq 0$, for, if $e_\alpha = 0$ for all α, we have $l_{n-m+1,\beta} = 0$, $\beta = 1, \cdots, n - m$, that is, $w'_{n-m+1} = y_{n-m+1}$ and $w_{n-m+1} = x_{n-m+1}$, hence x and y have the same distribution. Assuming $e_1 \neq 0$, set $t_\alpha = -e_\alpha t_{n-m+1}$, $\alpha = 2, \cdots, n - m$, in (14), obtaining

$$f(t_1) + \sum_{\alpha=2}^{n-m} f(-e_\alpha t_{n-m+1}) + f(t_{n-m+1}) + \sum_{\beta=1}^{n-m} f(-l_{1\beta}(t_1 + e_1 t_{n-m+1})) = 0, \tag{15}$$

now, recalling that $f(0) = 0$, set $t_{n-m+1} = 0$, getting $f(t_1) + \sum_{\beta=1}^{n-m} f(-l_{1\beta} t_1)$. Evaluating this with argument $t_1 + e_1 t_{n-m+1}$, and substituting back in (15) it appears that

$$f(t_1) + f(t_{n-m+1}) + \sum_{\alpha=2}^{n-m} f(-e_\alpha t_{n-m+1}) = f(t_1 + e_1 t_{n-m+1}). \tag{16}$$

Now setting $t_1 = 0$ in (16) we have the relation

$$f(t_{n-m+1}) + \sum_{\alpha=2}^{n-m} f(-e_\alpha t_{n-m+1}) = f(e_1 t_{n-m+1}).$$

thus we have finally $f(t_1) + f(e_1 t_{n-m+1}) = f(t_1 + e_1 t_{n-m+1})$, or, since $e_1 \neq 0$, $f(t_1 + t_2) = f(t_1) + f(t_2)$. The last relation implies that $f(t) = ct$, since $f(t)$ is continuous. Now replace $f(t)$ by ct in (13), getting $c\left\{\sum_{\alpha=1}^{n-m+1} t_\alpha - \sum_{\alpha=1}^{n-m+1}\sum_{\beta=1}^{n-m} l_{\alpha\beta} t_\alpha\right\} = 0$, that is, either $c = 0$, or $\sum_{\beta=1}^{n-m} l_{\alpha\beta} = 1$ for all α. We conclude then that $\varphi(t) = \psi(t)$, unless $\sum_\beta l_{\alpha\beta} = 1$ for all α. If $\sum_\beta l_{\alpha\beta} = 1$ for all α we have $\varphi(t) = e^{ct}\psi(t)$. $\varphi(-t) = \overline{\varphi(t)}$ and $\psi(-t) = \overline{\psi(t)}$, hence c is of the form $c = ia$, where a is real, in other words $\varphi(t) = e^{iat}\psi(t)$, thus concluding the proof of the theorem.

It was assumed in Theorem IVa that $\varphi(t)$ has no zeros. If $\varphi(t)$ has zeros we have proved that, for an interval $|t| < \epsilon$, $\varphi(t) = \psi(t)$ (or $\varphi(t) = e^{iat}\psi(t)$). This does not necessarily imply the result of Theorem IVa, but it does imply at least that if the kth moments of x and of y (or of $y - a$) both exist they are equal.

The last result in this series can be proved by methods similar to those used in Theorem IVa.

THEOREM IVB: *Let x and y satisfy the hypotheses of Theorem* IVa. *Suppose, moreover, that $m > 2(n - m)$, that the rank of $\| l_{\alpha\beta} \|$ is $n - m$, and that $\sum_{\beta=1}^{n-m} l_{\alpha\beta} \neq 1$ for at least $2m - n$ values of α. Then, if there exist constants $\{c_\alpha\}$ such that the set $\{c_\alpha w'_\alpha\}$ have the same joint distribution as $\{w_\alpha\}$, it follows that, for some α, $c_\alpha y$ has the same distribution as x.*

3. Application to Composite Hypotheses. The results of section 2 have a significant application in the theory of testing composite hypotheses. Suppose that x has a distribution of the form $F(x, \theta_1, \theta_2)$, and that the hypothesis $\theta_2 = \theta_2^0$ is to be tested, without reference to the value of θ_1. We assume that the parameters are independent, i.e., $F(x, \theta_1, \theta_2) \equiv F(x, \theta_1', \theta_2')$ implies that $\theta_1 = \theta_1'$ and $\theta_2 = \theta_2'$. It is true in a wide class of important cases that, given a sample $x_1, \cdots, x_n$ from the distribution $F(x, \theta_1, \theta_2)$, there exist functions $y_\alpha(x_1, \cdots, x_n)$, $\alpha = 1, 2, \cdots, p$, such that $\{y_\alpha\}$ have joint distribution independent of θ_1, but depending on θ_2. Now if the $\{y_\alpha\}$ are such that their joint distribution redetermines the original distribution, except for θ_1, one can reasonably use the p-dimensional distribution of the $\{y_\alpha\}$ for testing the hypothesis $\theta_2 = \theta_2^0$, thus reducing the composite hypothesis to a simple hypothesis. In testing this simple hypothesis, every alternative hypothesis (corresponding to a value of θ_2) determines a distribution of x among the alternatives $F(x, \theta_1, \theta_2)$ except for the unknown θ_1, that is, there is a one-to-one correspondence between the two sets of alternative hypotheses, expressed by the fact that if $\theta_2' = \theta_2''$ then the distributions of the set $\{y_\alpha\}$ corresponding to $\theta_2 = \theta_2'$ and $\theta_2 = \theta_2''$ must be different.

Suppose, for example, that it is desired to test whether $y = x - a$ for some a has the distribution $F(y, \theta^0)$, with the assumption that, for some a, y has the distribution $F(y, \theta)$. Given a sample one can form the set $w_\alpha = x_\alpha - x_n$, $\alpha = 1, 2, \cdots, n - 1$, obtaining the distribution $G(w_1, \cdots, w_{n-1}, \theta)$; now consider the simple hypothesis $\theta = \theta^0$, knowing that G determines θ, by Theorem Ia. Similarly one can test whether cx, for some $c \neq 0$, has distribution $F(y, \theta^0)$, by forming $w_\alpha = x_\alpha/x_n$, $\alpha = 1, \cdots, n - 1$, or by expressing $(x_1, \cdots, x_n)$ in spherical coordinates and considering the angular coordinates, according to whether both positive and negative or only positive values of c are to be allowed.

In the same way one can test the hypothesis $\theta = \theta^0$ under the assumption that $c(x - a)$ has distribution $F(y, \theta)$ by forming $w_\alpha = \dfrac{x_\alpha - x_n}{x_{n-1} - x_n}$, $\alpha = 1, \cdots, n - 2$, or by expressing $(x_1 - x_n, \cdots, x_{n-1} - x_n)$ in spherical coordinates and considering the angular coordinates.

Theorem IVa may be applied to analogous problems, in which the hypothesis $\theta = \theta^0$ is to be tested under the assumption that $y = u - \Sigma a_i x_i$ has distribution $F(y, \theta)$ for fixed values of the x_i, with the a_i unknown. In such problems there exist linear combinations of the observed values of y which are independent of the a_i. By Theorem IVa, under certain conditions the joint distribution of these linear combinations determines the original distribution of y, without regard to the a_i

In applying some of the preceding results we must verify in certain cases that the zeros of $\int e^{itx}\, dF(x)$ are nowhere dense, for a certain distribution function. By a change of variable the condition of Theorem Ib can be stated in this form; moreover if $F(x)$ satisfies this condition it is evident that it satisfies the condition of Theorem Ia. A sufficient condition applicable to a considerable class of cases has been obtained by Levinson [4]; if $f(x)$ is $O(e^{-\theta(x)})$ as $x \to \infty$, where $\theta(x)$ is monotone and $\int_1^\infty \dfrac{\theta(x)}{x^2}\, dx$ diverges to ∞, then $\int e^{itx} f(x)\, dx$ cannot vanish on an interval without vanishing identically. It is evident that it is likewise sufficient if the corresponding condition holds as $x \to -\infty$ instead of $+\infty$. In particular, if there exists A such that $f(x) = 0$ for $x > A$ (or for $x < A$) it is a consequence of the Levinson result that $\int e^{itx} f(x)\, dx$ has no intervals of zeros. It can be established easily that if $f(x)$ is majorized by $|x|^{-(1-\epsilon)}$, $\epsilon > 0$, in the neighborhood of the origin, then $\int e^{it \log |x|} f(x)\, dx$ has no intervals of zeros.

As a simple example consider the rectangular distribution on (0, 1). Let $(x - a)/r$ have this distribution with a unknown, $r > 0$, and suppose that we are interested only in r. Given a sample $x_1, \cdots, x_n$ form the functions $y_\alpha = (x_\alpha - x_n)/r$, $\alpha = 1, \cdots, n - 1$. Set $y_M = \max(y_\alpha, 0)$, $y_L = \min(y_\alpha, 0)$. Then it can be shown that $y_1, \cdots, y_{n-1}$ have probability density $(1 - y_M + y_L)$ in the region $-1 \leq y_\alpha \leq 1$, $y_M - y_L \leq 1$, zero elsewhere. $\psi = y_M - y_L$ is of course the quotient of the sample range by r. It can be shown that ψ has density $n(n - 1)(1 - \psi)\psi^{n-2}\,d\psi$. Theorem Ia makes it possible to base any tests not involving a on the distribution of the y_α, since if the y_α have the stated distribution then $(x - a)/r$ for some a must have the rectangular distribution.

Similarly, suppose $y = (x - a)/r$ has the distribution e^{-y}, $y > 0$, for some a, r. Then $w_\alpha = \dfrac{x_\alpha - x_n}{r}$, $\alpha = 1, 2, \cdots n - 1$, have distribution density $\dfrac{1}{n} e^{-\Sigma w_\alpha + n w_L}$, where $w_L = \min(0, w_\alpha)$. Again, the latter distribution may be used to estimate r.

Let us examine the distributions of functions of the type considered, in the case of normality. Assume that $x_1, \cdots, x_n$ are a sample of n observations from a normal distribution with unit variance and unknown mean. The variables $y_\alpha = x_\alpha - x_1$, $\alpha = 2, \cdots, n$, have a joint normal distribution with zero means and matrix of variances and covariances $\| A^{ij} \| = \| 1 + \delta_{ij} \|$. Then Theorem Ia shows that if $\{y_\alpha\}$ have this joint distribution then x is normally distributed with unit variance. Note that $\chi^2_{n-1} \equiv \Sigma A_{ij} y_i y_j \equiv \Sigma(x_\alpha - \bar{x})^2$. If we had $x = x'/\sigma$, then $\Sigma(x'_\alpha - \bar{x}')^2 = \sigma^2 x^2_{n-1}$, giving the estimate $\dfrac{1}{n-1} \Sigma(x'_\alpha - \bar{x}')^2$ for σ^2.

There are, of course, many ways in which the matrix $\| A_{ij} \|$ may be transformed into a diagonal matrix in order to obtain a new set of independently distributed variates; one convenient set is the set $\sqrt{\tfrac{1}{2}}\, y_2$, $\sqrt{\tfrac{2}{3}}\,(y_3 - \tfrac{1}{2}y_2), \cdots,$ $\sqrt{\dfrac{n-1}{n}}\left(y_n - \dfrac{1}{n-1}\sum_{\alpha=2}^{n-1} y_\alpha\right)$. In terms of the original x's we have $\sqrt{\tfrac{1}{2}}\,(x_2 - x_1)$ $\sqrt{\tfrac{2}{3}}\,(x_3 - \tfrac{1}{2}(x_1 + x_2))$, $\sqrt{\dfrac{n-1}{n}}\left(x_n - \dfrac{1}{n-1}\sum_{\alpha=1}^{n-1} x_\alpha\right)$; these functions of the data are independently distributed according to the normal distribution with zero mean and unit variance.

Similarly, in the case of a sample $x_1, \cdots, x_n$ from a normal distribution with zero mean and unknown variance, there exists a set of $n - 1$ functions with distributions independent of the variance. A convenient set of functions is the set

$$t_m = \frac{\sqrt{m}\, x_{m+1}}{\sqrt{\sum_{i=1}^{m} x_i^2}}; \qquad m = 1, \cdots, n - 1.$$

It is known (see Bartlett [1]) that the variables t_m are independently distributed according to student t-distributions with m degrees of freedom respectively. The set t_m determines the set of angular coordinates obtained by expressing $x_1, \cdots, x_n$ in spherical coordinates, hence we can conclude, conversely, that if $\{t_m\}$ have this joint distribution then x is normal with mean zero.

Finally we can eliminate both mean and variance. Suppose $x_1, \cdots, x_n$ are a sample from some normal distribution. The variables

$$u_m = \sqrt{\frac{m}{m+1}}\left\{x_{m+1} - \frac{1}{m}\sum_{i=1}^{m} x_i\right\}, \qquad m = 1, 2, \cdots, n-1,$$

are normal and independent with mean zero and some variance. Then we have the set

$$t_r' = \frac{\sqrt{r\left(\frac{r+1}{r+2}\right)}\left\{x_{r+2} - \frac{1}{r+1}\sum_{i=1}^{r+1} x_i\right\}}{\sqrt{\sum_{j=1}^{r}\frac{j}{j+1}\left\{x_{j+1} - \frac{1}{j}\sum_{i=1}^{j} x_i\right\}^2}}, \qquad r = 1, \cdots, n-2,$$

independently distributed according to t-distributions with r degrees of freedom respectively. It may be convenient for computational purposes to make use of the identity

$$\sum_{j=1}^{r}\frac{j}{j+1}\left\{x_{j+1} - \frac{1}{j}\sum_{i=1}^{j} x_i\right\}^2 \equiv \sum_{j=1}^{r+1}\left(x_j - \frac{1}{r+1}\sum_{i=1}^{r+1} x_i\right)^2 \equiv \sum_{j=1}^{r+1}(x_j - \bar{x}_{(r+1)})^2.$$

We then have

$$t_r' = \frac{\sqrt{r\left(\frac{r+1}{r+2}\right)}(x_{r+2} - \bar{x}_{(r+1)})}{\sqrt{\sum_{i=1}^{r+1}(x_i - \bar{x}_{(r+1)})^2}}, \qquad r = 1, \cdots, n-2.$$

Now, by Theorem IIIc, we know that if the set $\{t_r'\}$ has this specified distribution then x must be distributed according to some normal distribution. The set $\{t_r'\}$ may be used to test the goodness of fit of the observations to normality, by first adjusting the set $\{t_r'\}$ to a standard basis of comparison, i.e., by considering $F_r(t_r')$, where F_r is the corresponding cumulative distribution function and then applying, for example, a χ^2 goodness of fit test to these $n - 2$ quantities, with respect to the rectangular distribution on (0, 1).

REFERENCES

[1] M. S. Bartlett, *Proc. Camb. Phil Soc.* Vol. 30 (1934), pp. 164–169.
[2] J. Neyman and E. S. Pearson, *Biometrika*, Vol. XXA, (1928) pp. 175–240, 264–294.
[3] R. A. Fisher, *Proc. Roy. Soc.* A, Vol. 144 (1934), pp. 285–307.
[4] N. Levinson, *Proc. London Math. Soc.* Vol. 41 (1936), pp. 393–407.
[5] E. J. G. Pitman, *Biometrika* Vol. 30 (1939), pp. 391–421.

Princeton University,
Princeton, N. J.