The Annals of Statistics
1989, Vol. 17, No. 1, 195–208

ON THE STRUCTURE OF TRANSFORMATION MODELS

BY NIELS CHRISTIAN BANG JESPERSEN

Institute of Mathematical Statistics, University of Copenhagen

In this paper we characterize transformation models by means of the functional form of the densities. We discuss sufficiency of the pair (t, π) where t is an equivariant estimator and π is a maximal invariant. Furthermore, we introduce and discuss the algebraic concept of structural sufficiency. This gives rise to an example of a transformation model where (t, π) is nonsufficient.

1. Introduction. In the analysis of statistical models it is sometimes convenient to make use of invariance properties of the model in question. For instance, the *invariance principle* (see Lehmann [20] or Hall, Wijsman and Ghosh [16]) is a widely accepted and frequently used statistical tool. Closely related to this concept is the notion of transformation models. Let E be a sample space, Θ a parameter set and G a group acting on E and Θ. In our setup a transformation model is a family of probability measures $(P_\vartheta)_{\vartheta \in \Theta}$ with the property

$$\forall \vartheta \in \Theta \quad \forall g \in G: P_{g\vartheta}(A) = P_\vartheta(g^{-1}A) \tag{1.1}$$

for measurable sets A.

Though much attention has been given to the study of particular transformation models (see, e.g., Andersson, Brøns and Jensen [5], Andersson and Perlman [4], Eriksen [14] or Jensen [18]) a more general treatment of transformation models has only been given in some special cases (see, e.g., Barndorff-Nielsen, Blæsild, Jensen and Jørgensen [8], Eaton [12], Eriksen [13], Fraser [15], Roy [22] and Rukhin [23]) using different setups. The aim of this paper is to introduce a basic setup for general transformation models. In this setup we will characterize the models (1.1) by means of their densities. Furthermore we will discuss the concept of unique maximum likelihood estimation. If t: $E \to \Theta$ is a MLE and π is a maximal invariant it is sometimes assumed that (t, π) is sufficient (see, e.g., Barndorff-Nielsen [6] and [7] and Barndorff-Nielsen, Blæsild, Jensen and Jørgensen [8]). We will give conditions ensuring (t, π) to be sufficient and, by a nontrivial example, show that (t, π) is indeed not always sufficient.

In this paper we will make some apparently harmless topological regularity assumptions. These assumptions are nevertheless strong enough to imply that the results, proofs, etc., almost only depend on the algebraic structure of the groups and actions involved. We will rely heavily on the theory of invariant measures and group theory at a fairly elementary level. For an extensive exposition of the theory of invariant measures see Bourbaki [10] or Reiter [21].

Received February 1986; revised March 1988.
AMS 1980 *subject classifications*. Primary 62A05, 62B05; secondary 57S99.
Key words and phrases. Transformation models, invariance, equivariant mappings, structural sufficiency, proper action, homogeneous space.

For a more introductory exposition see Andersson [2]. In the theory of invariant measures the notion of a proper action appears naturally. For more comments on proper actions see Andersson [3] and Wijsman [25].

It should be stressed that readers not familiar with topology and the theory of invariant measures can read the paper disregarding the details and still be able to understand the main results.

2. Characterization of transformation models. At first, we will state the basic assumptions used throughout this paper and introduce some notation. A locally compact topological space (group) with a denumerable basis for the topology is usually called a *LCD space* (*group*). A LCD space is in fact a locally compact Polish space so it is indeed σ-compact, metrizable with a complete metric and there exists a countable dense subset.

Let E denote the sample space, Θ the parameter set and G a group. We will assume that G is a LCD group acting continuously on the LCD spaces E and Θ (both actions being left actions). The action on E induces an action on the set of probability measures on E by $gP(A) = P(g^{-1}A)$. Let $\pi: E \to G \setminus E$, $\tilde{\pi}: E \times \Theta \to G \setminus E \times \Theta$ denote the orbit projections under G's action on E, respectively, G's diagonal action on $E \times \Theta$, i.e., $g(x, \vartheta) = (gx, g\vartheta)$. The orbit space $G \setminus E$ is the set of orbits $Gx, x \in E$ and correspondingly $G \setminus E \times \Theta$ is the set of orbits $G(x, \vartheta), (x, \vartheta) \in E \times \Theta$ under the action of G on $E \times \Theta$. We equip the orbit spaces with the finest topology making π, respectively, $\tilde{\pi}$ continuous. Sometimes we will assume that G acts properly on Θ, i.e., the inverse of compact sets under the mapping $(g, \vartheta) \to (\vartheta, g\vartheta)$ are compacts too. It is worth noting that the orbit space in this case is a LCD space as well (see, e.g., Bourbaki [9] or [10]).

We will restrict the attention to statistical models $P = (P_\vartheta)_{\vartheta \in \Theta}$ which are dominated by a measure μ, being relatively invariant with multiplier $\chi: G \to \mathbb{R}_+$, i.e., $\forall g \in G$: $\mu(gA) = \chi(g)\mu(A)$ for measurable sets A. For convenience we will assume that μ has support E. We will assume that the densities $f_\vartheta(x)$ are jointly continuous in ϑ and x.

In the paper we will use the concept of a modulator $m: \Theta \to \mathbb{R}_+$ for the multiplier χ. A modulator is a continuous function with the property $\forall g \in G$, $\forall \vartheta \in \Theta$: $m(g\vartheta) = \chi(g)m(\vartheta)$. Note that if G acts properly on Θ, a modulator for χ does exist (see Bourbaki [10]).

First, we need an easy but fundamental lemma.

LEMMA 2.1. $(P_\vartheta)_{\vartheta \in \Theta}$, $P_\vartheta = f_\vartheta \mu$, *is a transformation model if and only if*

$$\forall \vartheta \in \Theta, \forall g \in G, \forall x \in E: \qquad f_\vartheta(x) = f_{g\vartheta}(gx)\chi(g). \tag{2.1}$$

PROOF. gP_ϑ has density $f_\vartheta(g^{-1}x)$ w.r.t. $g\mu = \chi(g)^{-1}\mu$, i.e., $f_\vartheta(g^{-1}x)\chi(g)^{-1}$ w.r.t. μ. Now $(P_\vartheta)_{\vartheta \in \Theta}$ is a transformation model if and only if gP_ϑ equals $P_{g\vartheta}$, i.e., if and only if $f_{g\vartheta}(x) = f_\vartheta(g^{-1}x)\chi(g)^{-1}$ which is equivalent to (2.1). □

The following theorem gives the basic structure of transformation models.

THEOREM 2.1. *If* $(P_\vartheta)_{\vartheta\in\Theta}$, $P_\vartheta = f_\vartheta\mu$, *is a transformation model and* $m: \Theta \to \mathbb{R}_+$ *is a modulator for* χ, *then there exists a continuous function* $p: G \setminus E \times \Theta \to \mathbb{R}_+$ *with*

$$\forall\, \vartheta \in \Theta, \forall\, x \in E: \qquad f_\vartheta(x) = p(\tilde{\pi}(x, \vartheta))/m(\vartheta), \tag{2.2}$$

$$\forall\, \vartheta \in \Theta: \qquad \int_E p(\tilde{\pi}(x, \vartheta))\, d\mu(x) = m(\vartheta). \tag{2.3}$$

On the other hand, if $p: G \setminus E \times \Theta \to \mathbb{R}_+$ *is a continuous function so that*

$$\forall\, \vartheta \in \Theta: \qquad \int_E p(\tilde{\pi}(x, \vartheta))\, d\mu(x) < +\infty \tag{2.4}$$

and $m(\vartheta) = \int_E p(\tilde{\pi}(x, \vartheta))\, d\mu(x)$ *is continuous, then it is in fact a modulator and* (2.2) *defines the densities of a transformation model.*

PROOF. Let $(P_\vartheta)_{\vartheta\in\Theta}$ be a transformation model and $m: \Theta \to \mathbb{R}_+$ a modulator for χ. Lemma 2.1 shows that $f_\vartheta(x) = f_{g\vartheta}(gx)\chi(g) = f_{g\vartheta}(gx)m(g\vartheta)/m(\vartheta)$ so the mapping $\psi: E \times \Theta \to \mathbb{R}_+$ defined by $\psi(x, \vartheta) = f_\vartheta(x)m(\vartheta)$ is invariant under the diagonal action of G on $E \times \Theta$, i.e., ψ factorizes through the orbit projection $\tilde{\pi}$, $\psi = p \circ \tilde{\pi}$, where $p: G \setminus E \times \Theta \to \mathbb{R}_+$ is continuous. This establishes (2.2) and (2.3) is trivial. On the other hand, if $p: G \setminus E \times \Theta \to \mathbb{R}_+$ is a continuous function fulfilling (2.4), then $m: \Theta \to \mathbb{R}_+$ defined by $m(\vartheta) = \int_E p(\tilde{\pi}(x, \vartheta))\, d\mu(x)$ is continuous, by definition, and

$$\begin{aligned} m(g\vartheta) &= \int_E p(\tilde{\pi}(x, g\vartheta))\, d\mu(x) = \int_E p(\tilde{\pi}(g^{-1}x, \vartheta))\, d\mu(x) \\ &= \int_E p(\tilde{\pi}(x, \vartheta))\, dg^{-1}\mu(x) = \int_E p(\tilde{\pi}(x, \vartheta))\chi(g)\, d\mu(x) \\ &= \chi(g)m(\vartheta), \end{aligned}$$

showing that m is a modulator. Furthermore f_ϑ's defined by (2.2) obviously satisfies (2.1) and, by assumption, the f_ϑ's are densities. □

REMARK. If G acts transitively and properly on Θ, then m, as defined in the second half of the proposition, is automatically continuous.

DEFINITION 2.1. If $(P_\vartheta)_{\vartheta\in\Theta}$ is a transformation model we will denote by p *the associated model function.*

If G acts properly on Θ (2.3) and (2.4) can be formulated in a more natural way. If G acts properly on Θ the diagonal action of G on $E \times \Theta$ is proper as well (see Bourbaki [9], Chapter 3, Section 4, Exercise 10c), i.e., $G \setminus E \times \Theta$ is locally compact and the orbit projection $\tilde{\pi}: E \times \Theta \to G \setminus E \times \Theta$ is proper. $\tilde{\pi}_\vartheta: E \to G \setminus E \times \Theta$ defined by $\tilde{\pi}_\vartheta(x) = \tilde{\pi}(x, \vartheta)$ is a composition of the two proper mappings $x \to (x, \vartheta)$ and $\tilde{\pi}$ and hence a proper mapping. Therefore $\tilde{\pi}_\vartheta(\mu)$

is a well-defined measure on $G \setminus E \times \Theta$ so (2.3) and (2.4) can be reformulated as

$$\text{(2.3}'\text{)} \qquad \forall\, \vartheta \in \Theta\colon \quad \int_{G \setminus E \times \Theta} p\, d\tilde{\pi}_{\vartheta}(\mu) = m(\vartheta),$$

$$\text{(2.4}'\text{)} \qquad \forall\, \vartheta \in \Theta\colon \quad \int_{G \setminus E \times \Theta} p\, d\tilde{\pi}_{\vartheta}(\mu) < +\infty.$$

This shows that if G acts properly on Θ transformation models can be constructed ad libitum by choosing model functions being $\tilde{\pi}_{\vartheta}(\mu)$-integrable . If $\tilde{\pi}_{\vartheta}(\mu)$ is not a well-defined Radon measure it sure is possible that no transformation model exist. Take, for instance, $\Theta = \{\vartheta_0\}$, the action on Θ being the trivial action ($\forall\, g \in G$: $g\vartheta_0 = \vartheta_0$). A transformation model, in this context, is simply an invariant probability measure. But an invariant measure does not always exist (see Example 2.1). When applying Theorem 2.1 the main obstacle is to identify the orbit space $G \setminus E \times \Theta$ and, in the case of a proper action, to identify the measures $\tilde{\pi}_{\vartheta}(\mu)$. These identifications are simplified when dealing with spaces having the following structure.

DEFINITION 2.2. E is a *TT-space* if E is homeomorphic to a product space $E_1 \times E_2$ so that G acts *trivially* on E_1, i.e., $\forall\, g \in G, \forall\, x_1 \in E_1$: $gx_1 = x_1$ and G acts *transitively* on E_2, i.e., $\forall\, x_2, y_2 \in E_2, \exists\, g \in G$: $x_2 = gy_2$.

REMARK. TT is an abbreviation for trivial transitive.

Note that if G acts transitively on E, then E is a TT-space (take E_1 degenerate). Likewise, if G acts freely on E, i.e., $gx = hx \Rightarrow g = h$, and there exists a homeomorphic orbit representation of $G \setminus E$, then E is a TT-space (take $E_1 = G \setminus E$, $E_2 = G$). The notion of TT-spaces covers thus the simple cases of transitive respective free actions.

TT-spaces have some nice properties. If $E \simeq E_1 \times E_2$ is a TT-space, then $G \setminus E$ is homeomorphic to E_1 and E_2 is a homogeneous space. The latter means that E_2 is homeomorphic to a quotient space G/K where K is a closed subgroup of G, take $K = G_{x_2} = \{g \in G | gx_2 = x_2\}$ for an $x_2 \in E_2$. Note that relatively invariant measures on homogeneous spaces are equivalent and relatively invariant measures with the same multiplier are proportional (see, e.g., Bourbaki [10]). If μ is a relatively invariant measure on E, then for $A \subseteq E_1$, $B \subseteq E_2$, measurable sets, $\mu(A \times gB) = \mu(gA \times gB) = \mu(g(A \times B)) = \chi(g)\mu(A \times B)$ showing that the measures on E_2 ν_A: $B \to \mu(A \times B)$ are relatively invariant with the same multiplier χ and hence proportional, i.e., $\mu(A \times B) = \kappa(A)\nu(B)$, where $\kappa(A)$ is a proportionality factor and ν is a relatively invariant measure with multiplier χ. It is easily shown that κ is a measure and hence $\mu = \kappa \otimes \nu$. Noting that the measures κ and μ are determined uniquely up to a norming factor we see that the relatively invariant measures on E have a very simple form.

Finally, G acts properly on E if and only if G acts properly on $E_2 \simeq G/K$. Now, it is easily shown that G acts properly on G/K if and only if K is compact so G acts properly on E if and only if the isotropic groups are compacts. The "only if" part of this statement is trivial whereas the "if" part is a special and very useful feature for TT-spaces.

PROPOSITION 2.1. *Let $E = E_1 \times E_2$ be a TT-space and assume that G acts transitively on Θ. Fix $\vartheta_0 \in \Theta$ and let $L = G_{\vartheta_0} = \{g \in G | g\vartheta_0 = \vartheta_0\}$, the isotropic group of ϑ_0. $G \setminus E \times \Theta$ is homeomorphic to $L \setminus E$ $(= E_1 \times L \setminus E_2)$. ($L \setminus E$ denotes the orbit space under L's action on E.)*

PROOF. Since E_2 and Θ are homogeneous spaces it is obviously enough to show that, say, $G \setminus G/K \times G/L$ is homeomorphic to $L \setminus G/K$ where K and L are subgroups of G. Define

$$(2.5) \qquad \psi: G/K \times G/L \to L \setminus G/K, \qquad \psi(gK, \tilde{g}L) = L\tilde{g}^{-1}gK,$$

ψ is easily seen to be well-defined, invariant, onto and continuous (using the relevant quotient topologies). To see that ψ is maximal invariant let $g, \tilde{g}, h, \tilde{h} \in G$ with $\psi(gK, \tilde{g}L) = \psi(hK, \tilde{h}L)$. Then $L\tilde{g}^{-1}gK = L\tilde{h}^{-1}hK \Leftrightarrow \tilde{g}^{-1}g \in L\tilde{h}^{-1}hK$, i.e., $\exists\ l \in L$, $\exists\ k \in K$, with $\tilde{g}^{-1}g = l\tilde{h}^{-1}hk$. This implies $gK = \tilde{g}l\tilde{h}^{-1}hkK = \tilde{g}l\tilde{h}^{-1}hK$ and $\tilde{g}L = \tilde{g}l\tilde{h}^{-1}\tilde{h}l^{-1}L = \tilde{g}l\tilde{h}^{-1}\tilde{h}L$, showing that $(gK, \tilde{g}L) \sim_G (hK, \tilde{h}L)$ and hence that ψ is maximal invariant. To see that $L \setminus G/K$ and $G \setminus G/K \times G/L$ are homeomorphic it remains to show that the mapping $LgK \to \tilde{\pi}(gK, L)$ is continuous but this is trivial. □

REMARK. Fix $x_0 \in E$ and set $K = G_{x_0}$. By symmetry we have $G \setminus E \times \Theta \simeq E_1 \times K \setminus \Theta$. This can be a useful observation.

Now we can formulate Theorem 2.1 for TT-spaces.

THEOREM 2.2. *Assume that G acts properly and transitively on Θ and that $E \simeq E_1 \times E_2$ is a TT-space. Fix $\vartheta_0 \in \Theta$ and set $L = G_{\vartheta_0}$. If $(P_\vartheta)_{\vartheta \in \Theta}$, $P_\vartheta = f_\vartheta \mu$, $\mu = \kappa \otimes \nu$, is a transformation model, then there exists a continuous function p: $E_1 \times L \setminus E_2 \to \mathbb{R}_+$ with*

$$(2.6) \qquad \forall\, x_1 \in E_1, \forall\, x_2 \in E_2, \forall\, g \in G: \qquad f_{g\vartheta_0}(x_1, x_2) = p(x_1, Lg^{-1}x_2)/\chi(g),$$

$$(2.7) \qquad \int p\, d\kappa \otimes \tilde{\pi}_L(\nu) < +\infty.$$

($\tilde{\pi}_L$: $E_2 \to L \setminus E_2$ is the orbit projection.) On the other hand, if

$$p: E_1 \times L \setminus E_2 \to \mathbb{R}_+$$

is continuous satisfying (2.7), then, possibly after a normalization of p, (2.6) defines a transformation model.

REMARK. Note that L is compact so $m(g\vartheta_0) = \chi(g)$ is a well-defined continuous function and is easily seen to be a modulator for χ. This is the reason why χ and not m appears in the formula (2.6). Under the assumptions in the theorem one can construct transformation models ad libitum as soon as $L = G_{\vartheta_0}$, $L \setminus E_2$ and $\tilde{\pi}_L(\nu)$ have been identified.

We will comment a little bit on unique maximum likelihood estimation. If $(P_\vartheta)_{\vartheta \in \Theta}$, $P_\vartheta = f_\vartheta \mu$ is a transformation model admitting unique maximum likelihood estimation, then the maximum likelihood estimator (MLE) t: $E \to \Theta$ is well known to be equivariant, i.e., $\forall\, g \in G$, $\forall\, x \in E$: $t(gx) = gt(x)$. In this case $G_x \subseteq G_{t(x)}$. Therefore, if G acts properly on Θ the isotropic group $G_{t(x)}$ is compact and hence G_x is compact. If E is a TT-space we can then conclude that G acts properly on E and an invariant measure exists. It is thus no restriction to assume that $\mu = \kappa \otimes \nu$ is invariant. Fix $\tilde{x}_2 \in E_2$ and let $K = G_{\tilde{x}_2}$. According to Theorem 2.2 and the remark to Theorem 2.2 the densities have the form $f_\vartheta(x_1, g\tilde{x}_2) = p(x_1, Kg^{-1}\vartheta)$, where p: $E_1 \times K \setminus \Theta \to \mathbb{R}_+$ is continuous. We then get the following result.

PROPOSITION 2.2. *Assume that E is a TT-space and G acts transitively and properly on Θ. Let $(P_\vartheta)_{\vartheta \in \Theta}$, $P_\vartheta = f_\vartheta\ \kappa \otimes \nu$, be a transformation model with $f_\vartheta(x_1, g\tilde{x}_2) = p(x_1, Kg^{-1}\vartheta)$. $(P_\vartheta)_{\vartheta \in \Theta}$ admits unique maximum likelihood estimation if and only if for each $x_1 \in E_1$ the mapping $K\vartheta \to p(x_1, K\vartheta)$ has a unique maximum at, say, $K\tilde{\vartheta}(x_1)$ with $K\tilde{\vartheta}(x_1)$ degenerate, i.e., $K \subseteq G_{\tilde{\vartheta}(x_1)}$. In this case the MLE is given by $t(x_1, g\tilde{x}_2) = g\tilde{\vartheta}(x_1)$.*

PROOF. Straightforward. □

We will close this section with some applications of Theorem 2.2 and Proposition 2.2.

EXAMPLE 2.1 (Multivariate location and scale parameter models). Take $E = \mathbb{R}^d$, $\Theta = H^+(d) \times \mathbb{R}^d$ and $G = AG(d)$. $AG(d) = \{[A, \alpha] | A \in GL(d), \alpha \in \mathbb{R}^d\}$ is the affine group of order d and $H^+(d)$ is the set of positive definite $d \times d$ matrices. The composition rule in $AG(d)$ is defined as follows $[A, \alpha][B, \beta] = [AB, A\beta + \alpha]$, $[A, \alpha]^{-1} = [A^{-1}, -A^{-1}\alpha]$ the unity being $[I, 0]$. The actions are given by

$$\begin{aligned} &AG(d) \times \mathbb{R}^d \to \mathbb{R}^d, \\ &([A, \alpha], x) \to Ax + \alpha, \end{aligned} \tag{2.8}$$

$$\begin{aligned} &AG(d) \times \left(H^+(d) \times \mathbb{R}^d\right) \to H^+(d) \times \mathbb{R}^d, \\ &([A, \alpha], (\Sigma, \xi)) \to (A\Sigma A^*, A\xi + \alpha). \end{aligned} \tag{2.9}$$

$\vartheta = (\Sigma, \xi)$ should be thought of as the covariance and the mean, respectively. Both actions are transitive, (2.9) is proper, the isotropic groups being homeomorphic to $O(d)$ and hence compact, whereas (2.8) is nonproper, the isotropic groups being homeomorphic to $GL(d)$ and hence noncompact. There exists no invariant measure on $\mathbb{R}^d$ under $AG(d)$, but the Lebesgue measure μ is relatively invariant with multiplier $\chi(A, \alpha) = |\det(A)|$. Take $\vartheta_0 = (I, 0)$. We thus have to identify $L = G_{\vartheta_0}$, $L \setminus \mathbb{R}^d$, the mapping $(x, g\vartheta_0) \to Lg^{-1}x$ and finally the measure $\tilde{\pi}_L(\mu)$.

Now, $L \simeq O(d)$ and $O(d) \setminus \mathbb{R}^d$ is well known to be homeomorphic to $[0, +\infty[$ via the identification $Lx \simeq \|x\|^2 = x^*x$. Let $(\Sigma, \xi) = (AA^*, \xi) = [A, \xi](I, 0) \in \Theta$ then $L[A, \xi]^{-1}x \sim \|[A, \xi]^{-1}x\|^2 = \|A^{-1}x - A^{-1}\xi\|^2 = (x - \xi)^*(A^{-1})^*A^{-1}(x - \xi) = (x - \xi)^*\Sigma^{-1}(x - \xi)$. It thus remains to identify $\tilde{\pi}_L(\mu)$. Letting γ denote the left-translation on the group $(\mathbb{R}_+, \cdot)$ we see that $\gamma(s^{-1})\tilde{\pi}_L(\mu) = s^{d/2}\tilde{\pi}_L(\mu)$ so $\tilde{\pi}_L(\mu)$ is relatively invariant with multiplier $s \to s^{d/2}$ and hence having density $s^{d/2-1}$ w.r.t. Lebesgue measure on $\mathbb{R}_+$.

We can thus conclude that the transformation models on $\mathbb{R}^d$ with parameter set $H^+(d) \times \mathbb{R}^d$ are exactly those of the form $P_{\Sigma, \xi} = f_{\Sigma, \xi}\mu$, μ Lebesgue measure, where $f_{\Sigma, \xi}(x) = p((x - \xi)^*\Sigma^{-1}(x - \xi))/\det(\Sigma)^{1/2}$ and $p\colon [0, +\infty[\to \mathbb{R}_+$ is a continuous function with $\int_0^\infty p(s)s^{d/2-1}\,ds < +\infty$.

This is a well-known result (see, e.g., Kelker [19]) and distributions with densities of this form are called *elliptic distributions*. Note finally that if $(P_{\Sigma, \xi})$ is a statistical model parametrized by the covariance and the mean, then it is a transformation model under the affine group and hence an elliptic distribution. Conversely, it is possible to show that if $(P_{\Sigma, \xi})$ is a transformation model with finite expectation and covariance, then the expectation equals ξ and the covariance is proportional to Σ.

EXAMPLE 2.2. Take $E = \Theta = H^+(d)$ and $G = GL(d)$ the general linear group of order d. The action is given by

$$\begin{aligned} GL(d) \times H^+(d) &\to H^+(d), \\ (A, \Sigma) &\to A\Sigma A^*. \end{aligned} \tag{2.10}$$

This action is transitive and proper, the invariant measure on $H^+(d)$ has density $S \to (\det S)^{-(d+1)/2}$ w.r.t. Lebesgue measure on $H^+(d)$. We are thus covered by Theorem 2.2. Take $\vartheta_0 = I$, then $L = G_I = O(d)$ and $O(d) \setminus H^+(d)$ can be represented by $\Lambda_d = \{(\lambda_1, \dots, \lambda_d) \in \mathbb{R}^d | \lambda_1 \geq \cdots \geq \lambda_d > 0\}$ using the identification $O(d)S \simeq$ the vector of ordered eigenvalues of S (see, e.g., Bourbaki [9]). Let $\Sigma = AA^* \in H^+(d)$, then $O(d)A^{-1}S(A^{-1})^* \simeq$ the vector of ordered eigenvalues of S w.r.t. Σ which we will denote $E(S; \Sigma)$. According to Anderson ([1], Theorem 3.3), $\tilde{\pi}_L(\mu)$ has density w.r.t. Lebesgue measure on Λ_d and the density is given by

$$\delta(\lambda_1, \dots, \lambda_d) = \prod_{i=1}^{d} \lambda_i^{-(d+1)/2} \prod_{1 \leq i < j \leq d} (\lambda_i - \lambda_j).$$

We can thus conclude that the transformation models on $H^+(d)$ are those of the form $P_\Sigma = f_\Sigma \mu$, μ Lebesgue measure on $H^+(d)$, with $f_\Sigma(S) =$

$p(E(S;\Sigma))(\det S)^{-(d+1)/2}$, where $p\colon \Lambda_d \to \mathbb{R}_+$ is a continuous function with

$$\int_{\Lambda_d} p(\lambda_1,\dots,\lambda_d)\prod_{i=1}^{d}\lambda_i^{-(d+1)/2}\prod_{1\le i<j\le d}(\lambda_i-\lambda_j)\,d(\lambda_1,\dots,\lambda_d) < +\infty.$$

It can be shown that the only $O(d)$ orbits of $H^+(d)$ consisting of one point only are those corresponding to λI_d, $\lambda > 0$. This facts shows, according to Proposition 2.2, that $(P_\Sigma)_{\Sigma\in H^+(d)}$ admits unique maximum likelihood estimation if and only if the associated model function p has a unique maximum at a point of the form $(\lambda,\dots,\lambda)\in\Lambda_d$ and the MLE is then given by $t(S)=\lambda S$. Letting

$$p(\lambda_1,\dots,\lambda_d)=\prod_{i=1}^{d}\lambda_i^{m/2}e^{-\lambda_i/2},\qquad m\ge d,$$

we see that p has a unique maximum at $(m,\dots,m)$ and p satisfies the integrability condition so p is the associated model function of a transformation model with unique MLE $t(S)=mS$, namely, the d-dimensional Wishart distribution with m degrees of freedom and unknown parameter Σ.

EXAMPLE 2.3 (Transformation model on the unit hyperboloid). Let $\underline{\varphi} = \mathrm{diag}(1,-1,-1,\dots,-1)$ be a $d\times d$ matrix and let φ_d denote the corresponding bilinear form on R^d. The unit hyperboloid is defined as $H_d=\{(x_1,\dots,x_d)^*\in\mathbb{R}^d \mid x_1>0,\ \varphi_d(x,x)=1\}$ and the group of hyperbolic transformations is

$$SH_d=\left\{A\in GL(d) \mid a_{11}>0, \det(A)=1,\ A^*\underline{\varphi}_d A=\underline{\varphi}_d\right\}.$$

SH_d acts transitively and properly on H_d by

$$\begin{aligned} SH_d\times H_d &\to H_d,\\ (A,x) &\to Ax \text{ (matrix multiplication)}. \end{aligned}\tag{2.11}$$

(See Vilenkin [24] or Jensen [18].) The invariant measure μ is given by

$$\mu(C)=\lambda_d\Big(\Big\{\mathbf{x}\in\mathbb{R}^d \mid 0<\varphi_d(\mathbf{x},\mathbf{x})\le 1,\ x_1>0,\ x/\sqrt{\varphi_d(\mathbf{x},\mathbf{x})}\in C\Big\}\Big),$$

for C a compact subset of H_d.

We will consider transformation models with $E=\Theta=H_d$ and $G=SH_d$ for $d\ge 3$. The above considerations imply that we are covered by Theorem 2.2. Let $\vartheta_0=(1,0,\dots,0)^*\in H_d$. Then

$$L=\left\{\begin{pmatrix}1 & 0\\ 0 & A\end{pmatrix}\,\middle|\, A\in SO(d-1)\right\},$$

where $SO(d-1)$ is the special orthogonal group of order $d-1$. One can readily check that $\mathbf{x}\sim_L\mathbf{y}$ if and only if $x_1=y_1$ since $SO(d-1)$ acts transitively on every sphere in $\mathbb{R}^{d-1}$. Therefore $L\setminus H_d$ can be identified with $[1,+\infty[$ using the identification $L\mathbf{x}\sim x_1=\varphi_d(x,\vartheta_0)$. If $\vartheta=A\vartheta_0$ with $A\in SH_d$, then $LA^{-1}\mathbf{x}\sim\varphi_d(A^{-1}x,\vartheta_0)=(A^{-1}x)^*\underline{\varphi}_d\vartheta_0=x^*(A^{-1})^*\underline{\varphi}_d\vartheta_0=x^*\underline{\varphi}_dA\vartheta_0=x^*\underline{\varphi}_d\vartheta=\varphi_d(x,\vartheta)$ which shows that the transformation models have the form, $P_\vartheta=f_\vartheta\mu$, $f_\vartheta(x)=p(\varphi(x,\vartheta))$, where $p\colon[1,+\infty[\to\mathbb{R}_+$ is a continuous function.

We will now identify the measure $\tilde{\pi}_L(\mu)$. For $t > 1$ we find, using Fubini's theorem,

$$\begin{aligned}
&\tilde{\pi}_L(\mu)([1, t]) \\
&\quad = \lambda_d\Big(\big\{x \in \mathbb{R}^d | 0 < x_1^2 - \cdots - x_d^2 \le 1,\ x_1 > 0,\ x_1^2 \le t^2\big(x_1^2 - \cdots - x_d^2\big)\big\}\Big) \\
&\quad = \lambda_d\Big(\big\{x \in \mathbb{R}^d | x_1^2 - 1 \le x_2^2 + \cdots x_d^2 \le \big(1 - 1/t^2\big)x_1^2,\ x_1 > 0\big\}\Big) \\
&\quad = \int_0^1 \lambda_{d-1}\Big(\big\{x \in \mathbb{R}^{d-1} | 0 \le x_1^2 + \cdots + x_{d-1}^2 \le \big(1 - 1/t^2\big)y^2\big\}\Big)\, dy \\
&\qquad + \int_1^t \lambda_{d-1}\Big(\big\{x \in \mathbb{R}^{d-1} | y^2 - 1 \le x_1^2 + \cdots + x_{d-1}^2 \le \big(1 - 1/t^2\big)y^2\big\}\Big)\, dy \\
&\quad = c\left[\int_0^t \big(1 - 1/t^2\big)^{(d-1)/2} y^{d-1}\, dy - \int_1^t \big(y^2 - 1\big)^{(d-1)/2}\, dy\right],
\end{aligned}$$

where c is a constant depending on d. This shows that $\tilde{\pi}_L(\mu)$ has density with respect to Lebesgue measure on $[1, +\infty[$ given by

$$\delta(t) = (\partial/\partial t)\tilde{\pi}_L(\mu)([1, t]) = c\big(t^2 - 1\big)^{(d-3)/2}(d - 1)/d$$

and hence the associated model functions have to satisfy $\int_1^\infty p(s)s^{d-3}\, ds < +\infty$.

Finally, it can be shown that the only degenerate L-orbit in H_d is the one corresponding to ϑ_0 so according to Proposition 2.2 the transformation model admits unique maximum likelihood estimation if and only if the model function p has a unique maximum at 1.

3. Structural sufficiency. Let $(P_\vartheta)_{\vartheta \in \Theta}$ be a transformation model admitting unique maximum likelihood estimation, t: $E \to \Theta$. In this section we will discuss sufficiency of the pair (t, π): $E \to \Theta \times G \setminus E$. Assume that (t, π) is sufficient. For a moment we will ignore problems with null-sets, continuity, measurability, etc. According to Neyman's theorem the densities have the form $f_\vartheta(x) = a_\vartheta(t(x), \pi(x))b(x)$. Then

$$f_\vartheta(x) = \frac{f_\vartheta(x)}{f_{t(x)}(x)} f_{t(x)}(x) = \frac{a_\vartheta(t(x), \pi(x))}{a_{t(x)}(t(x), \pi(x))} f_{t(x)}(x). \tag{3.1}$$

Now,

$$\begin{aligned}
f_{t(gx)}(gx)m(t(gx)) &= f_{gt(x)}(gx)m(gt(x)) \\
&= f_{t(x)}(x)\chi(g)^{-1}\chi(g)m(t(x)) \\
&= f_{t(x)}(x)m(t(x)) \quad [\text{according to } (2.1)]
\end{aligned}$$

so $f_{t(x)}(x)$ is of the form $g(\pi(x))/m(t(x))$ which inserted in (3.1) gives

$$f_\vartheta(x) = \frac{a_\vartheta(t(x), \pi(x))}{a_{t(x)}(t(x), \pi(x))} \frac{g(\pi(x))}{m(t(x))}, \tag{3.2}$$

showing that the density is a function of (t, π). According to the structure theorem in Section 2 the density is of the form $f_\vartheta(x) = p(\tilde{\pi}(x, \vartheta))/m(\vartheta)$ too. To be sure that f_ϑ is a function of (t, π), irrespective of the choice of p, we should demand that $\tilde{\pi}(x, \vartheta)$ is a function of (t, π). This should motivate the following definition.

DEFINITION 3.1. Let t: $E \to \Theta$ be an equivariant mapping, π: $E \to G \setminus E$ the orbit projection. (t, π) is *structurally sufficient* if, for each $\vartheta \in \Theta$, the mapping $\tilde{\pi}_\vartheta$: $E \to G \setminus E \times \Theta$, $\tilde{\pi}_\vartheta(x) = \tilde{\pi}(x, \vartheta)$, is a function of (t, π), say, $\tilde{\pi}_\vartheta(x) = \rho_\vartheta(t(x), \pi(x))$.

REMARK. If (t, π) is structurally sufficient, it is in fact a sufficient reduction in all transformation models because the densities then have the form $f_\vartheta(x) = p(\rho_\vartheta(t(x), \pi(x)))/m(\vartheta)$.

We can give a simple necessary and sufficient condition for structural sufficiency.

PROPOSITION 3.1. *(t, π) is structurally sufficient if and only if*

$$\forall \vartheta \in \Theta, \forall x \in E: \quad G_{t(x)} \subseteq G_\vartheta G_x. \tag{3.3}$$

PROOF. (t, π) is structurally sufficient if and only if

$$\forall \vartheta \in \Theta, \forall g \in G, \forall x \in E: \quad t(gx) = t(x) \Rightarrow \tilde{\pi}_\vartheta(x) = \tilde{\pi}_\vartheta(gx)$$

$$\Leftrightarrow \quad \forall \vartheta \in \Theta, \forall g \in G, \forall x \in E: \quad g \in G_{t(x)} \Rightarrow [\exists\, h \in G: h\vartheta = \vartheta,\ hx = gx]$$

$$\Leftrightarrow \quad \forall \vartheta \in \Theta, \forall g \in G, \forall x \in E: \quad g \in G_{t(x)} \Rightarrow [\exists\, h \in G_\vartheta: h^{-1}g \in G_x]$$

$$\Leftrightarrow \quad \forall \vartheta \in \Theta, \forall g \in G, \forall x \in E: \quad g \in G_{t(x)} \Rightarrow g \in G_\vartheta G_x,$$

which is exactly (3.3). □

REMARK. t is equivariant so $G_x \subseteq G_{t(x)}$. (3.3) says that even though $G_{t(x)}$ is larger than G_x it should not be too large.

COROLLARY 3.1. *Assume that G acts transitively on Θ. If the isotropic groups of Θ are* normal, *then (t, π) is structurally sufficient. If G acts freely on E, i.e., the isotropic groups of E consist only of the neutral element, (t, π) is structurally sufficient if and only if the isotropic groups of Θ are normal.*

PROOF. If the G_ϑ's are normal they are all equal because $G_{g\vartheta} = gG_\vartheta g^{-1} = G_\vartheta$ so $G_{t(x)} = G_\vartheta \subseteq G_\vartheta G_x$. If G acts freely on E, then $G_x = \{e\}$ so the condition (3.3) is equivalent to $\forall g \in G, \forall \vartheta \in \Theta$: $gG_\vartheta g^{-1} \subseteq G_\vartheta$. □

We will now introduce (see, e.g., Barndorff-Nielsen [6], [7])

DEFINITION 3.2. E and Θ are of the same orbit type if the G_x's and G_ϑ's are conjugates of one another, i.e., $\forall x \in E, \forall \vartheta \in \Theta, \exists g \in G$: $G_x = gG_\vartheta g^{-1}$.

REMARK. If E is a TT-space and E_2 is isomorphic to Θ, then E and Θ are of the same orbit type.

In the rest of this section we will assume that E and Θ are of the same orbit type. In this case the concept of structural sufficiency turns out to be rather trivial.

PROPOSITION 3.2. *(t, π) is structurally sufficient if and only if (t, π) is 1–1 and onto.*

PROOF. If (t, π) is structurally sufficient and $x \in E$, then $G_x \subseteq G_{t(x)}$ and we can find ϑ with $G_\vartheta = G_x$. According to (3.3) $G_x \subseteq G_{t(x)} \subseteq G_\vartheta G_x = G_x$ so $G_x = G_{t(x)}$ showing that (t, π) is 1–1. (t, π) is obviously onto. □

The proof of the above proposition motivates Definition 3.3.

DEFINITION 3.3. A subgroup $H \subseteq G$ is regular if

$$\forall g \in G: \qquad H \subseteq gHg^{-1} \Rightarrow H = gHg^{-1}. \tag{3.4}$$

REMARK. If H is regular any conjugate group gHg^{-1} is regular.

We then obtain

PROPOSITION 3.3. *If the G_ϑ's are regular, then (t, π) is structurally sufficient, i.e., 1–1 and onto.*

PROOF. Let $x \in E$ and choose $\vartheta \in \Theta$ with $G_x = G_\vartheta$. Now $G_{t(x)}$ is of the form $gG_\vartheta g^{-1}$ so $G_\vartheta = G_x \subseteq G_{t(x)} = gG_\vartheta g^{-1}$ which by the regularity of G_ϑ implies $G_x = G_{t(x)}$. □

This suggests a study of the concept of regularity. It is obvious that a normal subgroup is regular. Similarly, a maximal compact subgroup is regular.

EXAMPLE 3.1. Consider Example 2.2. We then have $E = \Theta = H^+(d)$ so E and Θ are of the same orbit type. Now, $G_I = O(d)$ which is known to be maximally compact and hence regular and by Proposition 3.3 we see that t has to be 1–1 and onto. This is in accordance with Example 2.2 in which we argued that $t(S) = \lambda S$ for some $\lambda > 0$.

We will now state a widely applicable result.

PROPOSITION 3.4. *Every compact subgroup of a Lie group of nonzero dimension is regular.*

For the notion of Lie group see, e.g., Bourbaki [11] or Hochschild [17]. The proposition is an easy corollary of the following result.

LEMMA 3.1. *Let H be a compact Lie group. If φ: $H \to H$ is a continuous injective homomorphism, then φ is onto.*

PROOF. Let H_e denote the connected component containing e. H_e is a closed normal subgroup of H. Since φ is a continuous homomorphism $\varphi(H_e) \subseteq H_e$. Let $L(H_e)$ denote the Lie algebra associated with H_e. Then φ in a canonical way induces an algebra homomorphism $L(\varphi)$: $L(H_e) \to L(H_e)$. φ being 1–1 implies that $L(\varphi)$ is 1–1 (see Bourbaki [11], Chapter III, Section 6). $L(H_e)$ is finite-dimensional so $L(\varphi)$ is onto, i.e., $L(\varphi)(L(H_e)) = L(H_e)$. According to Bourbaki ([11], Chapter III, Section 6) we then have $H_e = \varphi(H_e)$. Since H is locally connected H_e is open so H being compact implies that H/H_e is finite. φ defines in a canonical way a mapping $\bar{\varphi}$: $H/H_e \to H/H_e$ by $\bar{\varphi}(hH_e) = \varphi(h)H_e$. $\bar{\varphi}$ is easily seen to be 1–1 so the finiteness of H/H_e then implies that $\bar{\varphi}$ is onto as well. Let $h \in H$, choose $\tilde{h} \in H$ with $\bar{\varphi}(\tilde{h}H_e) = hH_e$, i.e., $\varphi(\tilde{h})H_e = hH_e$. Choose now $k \in H_e$ with $\varphi(\tilde{h}) = hk$ and $\tilde{k} \in H_e$ with $\varphi(\tilde{k}) = k^{-1}$. Then $\varphi(\tilde{h}\tilde{k}) = hkk^{-1} = h$ showing that φ is onto. □

PROOF OF PROPOSITION 3.4. Assume that H is a compact subgroup of G with $gHg^{-1} \subseteq H$. Define φ: $H \to H$ by $\varphi(h) = ghg^{-1}$. Now, φ is a continuous injective homomorphism and H is a compact Lie group so by Lemma 3.1 we indeed have that φ is onto, i.e., $gHg^{-1} = \varphi(H) = H$. □

REMARK. It is not true that every closed subgroup of a Lie group of nonzero dimension is regular.

We finally state a result for TT-space.

PROPOSITION 3.5. *Let E be a TT-space. (t, π) is structurally sufficient for* all *equivariant mappings t: $E \to \Theta$ if and only if the G_ϑ's are regular.*

We will close this section with an example of a transformation model which admits unique maximum likelihood estimation t with (t, π) nonsufficient.

EXAMPLE 3.2. Introduce

$$E = \left\{\left((x_k)_{k=N+1}^{\infty}; N\right) | N \in \mathbb{Z},\, x_k \in \{0,1\},\, k = N+1, N+2, \dots\right\}.$$

We equip M with the topology making

$$\iota\colon \{0,1\}^{\mathbb{N}} \times \mathbb{Z} \to E, \qquad \iota\left((x_k)_{k=1}^{\infty}; N\right) = \left((x_{k-N})_{k=N+1}^{\infty}; N\right)$$

a homeomorphism. Let $G = \{[\varphi, a] | \varphi \in \{0,1\}^{\mathbb{Z}},\ a \in \mathbb{Z}\}$, G is the semiproduct of $\{0,1\}^{\mathbb{Z}}$ and $\mathbb{Z}$, with composition rule $[\varphi, a][\psi, b] = [\varphi(a\psi), a + b]$, where

$(a\psi)(k) = \psi(k-a)$ and $(\varphi\psi)(k) = \varphi(k)\psi(k)$, the unit is $(\mathbf{0}, 0)$ and the inverse is given by $[\varphi, a]^{-1} = [(-a)\varphi, -a]$.

G acts on E by

$$G \times E \to E,$$

(3.5) $$\left([\varphi, a], \left((x_k)_{k=N+1}^{\infty}; N\right)\right) \to \left((\varphi(k)x_{k-a})_{k=a+N+1}^{\infty}; a+N\right).$$

(3.5) is transitive and proper. The invariant measure on E is given by $\mu = (\otimes_{i=1}^{\infty}\mu_i) \otimes \tau$, where $\mu_i(\{0\}) = \mu_i(\{1\}) = \frac{1}{2}$ and τ is counting measure on $\mathbb{Z}$. Notice that G is an LCD group, E is an LCD space and the isotropic group for $(\mathbf{0}, 0)$ is $G_{(\mathbf{0},0)} = K = \{[\varphi, 0] | \varphi(k) = 0 \ \forall\ k > 0\}$ which is homeomorphic to $\{0, 1\}^{\mathbb{N}}$ and hence compact but it is nonregular. Define $\Theta = E$. We will introduce a transformation model on E with parameter set Θ. Let $(p_k)_{k=1}^{\infty}$, $p_k \in]0, 1[$ be known reals. For $\vartheta = ((\vartheta_k)_{k=M+1}^{\infty}; M) \in \Theta$ we define the conditional distribution of $(X_k)_{k=N+1}^{\infty}$ given N as follows. $X_{N+1}, X_{N+2}, \dots$ are independent.

If $M \le N$, then

$$X_{N+k} \sim \begin{cases} \text{bin}(1, p_k), & \text{if } \vartheta_{k+N} = 0, \\ \text{bin}(1, 1-p_k), & \text{if } \vartheta_{k+N} = 1. \end{cases}$$

If $M \ge N$, then

$$X_{N+1}, X_{N+2}, \dots, X_M \sim \text{bin}\left(1, \tfrac{1}{2}\right),$$

$$X_{M+k} \sim \begin{cases} \text{bin}(1, p_k), & \text{if } \vartheta_{k+M} = 0, \\ \text{bin}(1, 1-p_k), & \text{if } \vartheta_{k+M} = 1. \end{cases}$$

The marginal distribution of N has density $q(M - \cdot\,)$ w.r.t. counting measure on $\mathbb{Z}$.

If $\forall\ k \in \mathbb{N}$: $p_k < \frac{1}{2}$, $p_1 = \frac{1}{4}$ and $\sum_{k=1}^{\infty}(1 - 2p_k) < +\infty$ and, say $\forall\ k > 1$: $q(k) = 0$ and finally $\cdots < q(-1) < q(0) < \frac{2}{3}q(1)$, then the above probability distributions on E give rise to a transformation model with a unique maximum likelihood estimator t: $E \to \Theta$, $t((x_k)_{k=N+1}^{\infty}; N) = ((x_k)_{k=N+2}^{\infty}; N+1)$ which is nonsufficient (details are left to the reader). This is thus an example of a transformation model where E and Θ are of the same orbit type, the maximum likelihood estimator exists uniquely but (t, π) is nonsufficient. As pointed out above this relies on the fact that the isotropic groups of $E = \Theta$ are nonregular.

Acknowledgments. I am indebted to Søren T. Jensen for drawing my attention to the problems treated in this paper and for many helpful and stimulating discussions. I would also like to thank Esben Kehlet for providing me with a sketch of the proof of Lemma 3.1. Finally, I would like to thank the referees for useful suggestions that improved the readability of the paper.

REFERENCES

[1] ANDERSON, T. W. (1958). *An Introduction to Multivariate Statistical Analysis*. Wiley, New York.

[2] ANDERSSON, S. A. (1978). Invariant measures. Technical Report 129, Dept. Statistics, Stanford Univ.

[3] ANDERSSON, S. A. (1982). Distributions of maximal invariants, using quotient measures. *Ann. Statist.* **10** 955–961.

[4] ANDERSSON, S. A. and PERLMAN, M. D. (1984). Two testing problems relating the real and complex multivariate normal distribution. *J. Multivariate Anal.* **15** 21–51.

[5] ANDERSSON, S. A., BRØNS, H. K. and JENSEN, S. T. (1983). Distributions of eigenvalues in multivariate statistical analysis. *Ann. Statist.* **11** 392–415.

[6] BARNDORFF-NIELSEN, O. E. (1983). Parametric statistical models and inference: Some aspects. Manuscript for the Forum Lectures at 14th European Meeting of Statisticians, Wrocław. Unpublished.

[7] BARNDORFF-NIELSEN, O. E. (1983). On a formula for the distribution of the maximum likelihood estimator. *Biometrika* **70** 343–365.

[8] BARNDORFF-NIELSEN, O. E., BLÆSILD, P., JENSEN, J. L. & JØRGENSEN, B. (1982). Exponential transformation models. *Proc. Roy. Soc. London Ser. A* **379** 41–65.

[9] BOURBAKI, N. (1960). *Eléments de Mathématique. Topologie General.* Chapitres 3 à 4. Hermann, Paris.

[10] BOURBAKI, N. (1963). *Eléments de Mathématique. Integration.* Chapitres 7 à 8. Hermann, Paris.

[11] BOURBAKI, N. (1972). *Eléments de Mathématique. Groupes et Algebres de Lie.* Chapitres 2 à 3. Hermann, Paris.

[12] EATON, M. L. (1983). *Multivariate Statistics.* Wiley, New York.

[13] ERIKSEN, P. S. (1984). $(k, 1)$ exponential transformation models. *Scand. J. Statist.* **11** 129–146.

[14] ERIKSEN, P. S. (1987). Proportionality of k covariance matrices. *Ann. Statist.* **15** 732–748.

[15] FRASER, D. A. S. (1979). *Inference and Linear Models.* McGraw-Hill, New York.

[16] HALL, W. J., WIJSMAN, R. A. and GHOSH, J. K. (1965). The relationship between sufficiency and invariance with applications in sequential analysis. *Ann. Math. Statist.* **36** 575–614.

[17] HOCHSCHILD, G. (1965). *The Structure of Lie Groups.* Holden-Day, San Francisco.

[18] JENSEN, J. L. (1981). On the hyperboloid distribution. *Scand. J. Statist.* **8** 193–206.

[19] KELKER, D. (1970). Distribution theory of spherical distributions and a location-scale parameter generalization. *Sankhyā Ser. A* **32** 419–430.

[20] LEHMANN, E. L. (1959). *Testing Statistical Hypotheses.* Wiley, New York.

[21] REITER, H. (1968). *Classical Harmonic Analysis and Locally Compact Groups.* Clarendon Press, Oxford.

[22] ROY, K. K. (1975). Exponential families of densities on an analytic group and sufficient statistics. *Sankhyā Ser. A* **37** 82–92.

[23] RUKHIN, A. L. (1974). Characterizations of distributions by statistical properties on groups. In *Statistical Distributions in Scientific Work* (G. P. Patil, S. Kotz and J. K. Ord, eds.) **3** 149–161. Reidel, Dordrecht.

[24] VILENKIN, N. YA. (1968). *Special Functions and the Theory of Group Representation. Transl. Math. Monographs* **22.** Amer. Math. Soc., Providence, R.I.

[25] WIJSMAN, R. A. (1985). Proper action in steps, with application to density ratios of maximal invariants. *Ann. Statist.* **13** 395–402.

CODAN INSURANCE, LTD
CODAN LIV
1899 FREDERIKSBERG C
DENMARK