A SPECIAL GROUP STRUCTURE AND EQUIVARIANT ESTIMATION[1]

By Robert H. Berk

University of California, Berkeley and University of Michigan

0. Summary. A special group structure common to many statistical problems is presented. In the context of equivariant estimation, this leads to a characterization of equivariant estimators (Section 3). Under further conditions, measurable equivariant estimators are characterized (Section 4). Some examples are presented in Section 5 and Section 6 contains a brief discussion of related work in which the special group structure is useful.

1. Introduction. The purpose of this paper is to discuss and illustrate a group structure common to many statistical problems. This is done in the context of estimation when a natural requirement of equivariance (called invariance by many writers) can be imposed. As a consequence, the structure leads to a specific characterization of equivariant estimators (Sections 3 and 4). We begin with some preliminaries.

X is a random variable taking its values in the measurable space $(\mathfrak{X}, \mathfrak{A})$. $\mathcal{P}$ is the model: a class of probability distributions for X (i.e., probability measures on $\mathfrak{A}$) and $\vartheta:\mathcal{P} \rightarrow \Theta$ is a mapping of $\mathcal{P}$ onto an arbitrary set Θ. For $P \varepsilon \mathcal{P}$, $\vartheta(P)$ will be referred to as the parameter associated with P. X is observed. It is then desired to estimate the parameter associated with its distribution, the latter being assumed a member of $\mathcal{P}$, but otherwise unknown. I.e., one desires a mapping $\varphi:\mathfrak{X} \rightarrow \Theta$ that has certain reasonable properties. Often, Θ is endowed with a σ-field $\mathcal{B}$ and estimators are required to be measurable. G is a group of one-one bimeasurable transformations of $\mathfrak{X}$ onto itself that leaves $\mathcal{P}$ invariant. That is, if X has distribution $P \varepsilon \mathcal{P}$, then for all $g \varepsilon G$, Pg^{-1}, the distribution of gX is also in $\mathcal{P}$. Thus g acting on $\mathfrak{X}$ induces a transformation $\bar{g}$ of $\mathcal{P}$ onto itself. ($\bar{g}P(A) = P(g^{-1}A)$, for all $A \varepsilon \mathfrak{A}$). It is easily verified that the collection $\bar{G}$, of induced transformations is also a group and the correspondence $g \rightarrow \bar{g}$ is a homomorphism (see [14]). The following concepts are central to the discussion.

1. Definition. A mapping φ of $\mathfrak{X}$ onto a set $\mathcal{Y}$ is said to commute with G if the map φg is a function of φ. That is, if $x, x' \varepsilon \mathfrak{X}$ and $\varphi x = \varphi x'$, then $\varphi gx = \varphi gx'$ for all $g \varepsilon G$.

Remark. If φ is one-one, it automatically commutes with G. A map that commutes with G induces in a natural way a group of transformations on its range. More specifically:

2. Proposition. *If φ mapping $\mathfrak{X}$ onto $\mathcal{Y}$ commutes with G, then to $g \varepsilon G$ corresponds a mapping $g_\varphi:\mathcal{Y} \rightarrow \mathcal{Y}$ so that* (i) $G_\varphi = \{g_\varphi : g \varepsilon G\}$ *is a group and* $g \rightarrow g_\varphi$ *is a homomorphism,* (ii) *for all* $x \varepsilon \mathfrak{X}$, $g \varepsilon G$, $\varphi gx = g_\varphi \varphi x$.

Received 6 February 1967; revised 16 May 1967.

[1] This work has been facilitated by Grant MH11493-01 from the Public Health Service and Grant GP-6008 from the National Science Foundation.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.* www.jstor.org

PROOF. For $g \varepsilon G$, $y \varepsilon \mathcal{Y}$, let $g_\varphi y = \varphi g x$, where $y = gx$. This definition is unambiguous since $y = \varphi x = \varphi x' \Rightarrow \varphi g x = \varphi g x'$. Statement (i) is easily verified (see, e.g. [14], pp. 213–214) and (ii) holds by construction.

REMARK. If G_φ is trivial φ is said to be G-invariant. More generally, we define the concept of equivariance.

3. DEFINITION. Suppose to G acting on $\mathfrak{X}$ there corresponds a group G^* acting on a set $\mathcal{Y}$ ($g^* \varepsilon G^*$ denotes the element corresponding to $g \varepsilon G$). A mapping $\varphi: \mathfrak{X} \to \mathcal{Y}$ is (G, G^*)-equivariant if for all $x \varepsilon \mathfrak{X}$, $g \varepsilon G$, $\varphi g x = g^* \varphi x$.

If φ is equivariant, the correspondence between G and G^* acting on range φ is necessarily a homomorphism. The following easily proved statement connects the two notions introduced above.

4. PROPOSITION. *A mapping φ of $\mathfrak{X}$ onto $\mathcal{Y}$ is (G, G^*)-equivariant $\Leftrightarrow$ φ commutes with G and $G_\varphi = G^*$.*

An equivariant estimation problem can arise if under the structure discussed above, $\vartheta: \mathcal{P} \to \Theta$ commutes with $\bar{G}$. For then, letting G^* denote the group ϑ induces on Θ, ϑ is $(\bar{G}, G^*)$-equivariant. Under these circumstances, it seems reasonable to inquire about the possibility of finding estimators φ that are (G, G^*)-equivariant. The rationale, oft stated (see e.g., [2], p. 209) is that since $\mathcal{P}$ is $\bar{G}$-invariant, we may arbitrarily transform X to gX without destroying the model. Under such circumstances, using a given estimator φ, we may estimate $\vartheta(\bar{g}P) = g^*\vartheta(P)$ by $\varphi(gX)$. It seems inconsistent to have $(g^*)^{-1}\varphi(gX) \neq \varphi(X)$, for then $(g^*)^{-1}\varphi(gX)$ can be put forth as a competing estimator of ϑ. The undesirability of this lies in the arbitrariness of g.

Hence we consider the possibility of obtaining (G, G^*)-equivariant estimators $\varphi: \mathfrak{X} \to \Theta$. Without measurability requirements, a complete characterization of such estimators is easily given. We discuss this in Section 2. We introduce a special group structure in Section 3 and discuss the particularly simple form equivariant estimators then have. In Section 4, measurability requirements are considered. No considerations of optimality are presented here, but it is known that minimax considerations sometimes lead to equivariant estimators [2], [5], [6], [8], [11], [12], [13], [20]. On occasion, they are even admissible, but not always [3], [4], [5], [10], [18]. The basic concepts used here relating to transformation groups may be found in [14] and [16]. Throughout, if G_0 is a subgroup of G, G/G_0 denotes the collection of left cosets of G_0.

2. Characterization of equivariant functions. We take as given the domain $(\mathfrak{X}, G)$ and range (Θ, G^*) and seek to characterize the equivariant mappings $\varphi: \mathfrak{X} \to \Theta$. Note that the necessary homomorphic relation exists between G and G^*, since they are both connected by homomorphisms with $\bar{G}$. Choose $x \varepsilon \mathfrak{X}$ and let $G_x = \{g \varepsilon G: gx = x\}$ be the isotropy subgroup of G at x. If $\varphi: \mathfrak{X} \to \Theta$ is (G, G^*)-equivariant, then necessarily, for all $g \varepsilon G_x$, $\varphi x = \varphi g x = g^* \varphi x$. That is, φx must be a fixed point of $(G_x)^*$, the image of G_x in G^*: $(G_x)^* \varphi x = \varphi x$. Clearly this is only possible if $(G_x)^*$ has a fixed point; i.e., if $\Theta_x = \{\theta: (G_x)^* \theta = \theta\}$ is not empty. If Θ_x is not empty for some $x \varepsilon \mathfrak{X}$, this is true of every point in Gx, the

G-orbit containing x ($\Theta_{gx} = g^*\Theta_x$). That the Θ_x be non-void is necessary and sufficient for there to exist (G, G^*)-equivariant mappings, as is shown by the following:

1. THEOREM. *Given* ($\mathfrak{X}$, G) *and* (Θ, G^*) *as above, there exist* (G, G^*)*-equivariant mappings* $\varphi:\mathfrak{X} \to \Theta \Leftrightarrow$ *for every* $x \varepsilon \mathfrak{X}$, Θ_x *is not empty.*

REMARK. In view of the above discussion, it is necessary and sufficient that the condition in the theorem hold for one point in each G-orbit.

PROOF. The necessity is outlined above. When for all $x \varepsilon \mathfrak{X}$, Θ_x is not empty, equivariant mappings φ may be constructed orbit-by-orbit. Simply choose $x \varepsilon \mathfrak{X}$ and a value $\varphi x \varepsilon \Theta_x$. Then at $y = gx \varepsilon Gx$, let $\varphi y = g^* \varphi x \varepsilon \Theta_y$. It is clear that this yields all equivariant mappings.

The preceding proof shows that as G is transitive on an orbit, an equivariant map is determined there by its value at one point. Thus all equivariant maps of $\mathfrak{X}$ may be obtained by piecing together the simpler equivariant maps on the oribits: Let $\mathfrak{X}_\alpha$ denote an orbit, $x_\alpha \varepsilon \mathfrak{X}_\alpha$, a reference point, G_α, the isotropy subgroup of G at x_α and let $\Theta_\alpha = \{\theta:(G_\alpha)^*\theta = \theta\}$. An equivariant map $\varphi_\alpha:\mathfrak{X}_\alpha \to \Theta$ is determined by $\varphi_\alpha x_\alpha \varepsilon \Theta_\alpha: \varphi_\alpha g x_\alpha = g^* \varphi_\alpha x_\alpha$ and if φ is defined to be φ_α on $\mathfrak{X}_\alpha$, φ is equivariant. Conversely, an equivariant φ is determined by $\{\varphi_\alpha\}$, its restrictions to the orbits. The collection $\{x_\alpha\}$ is called a cross-section; a set that intersects every orbit once. Thus an equivariant map is determined by its values on (any) cross-section. We shall see in the next section that in some cases, a very natural cross-section is available. (See also [21]).

The above construction is not always satisfactory if there are uncountably many orbits. For it is often desired to obtain measurable equivariant mappings and it is not clear how to piece together uncountably many maps and have the result be measurable. Of course if there are at most countably many orbits, then φ will be measurable $\Leftrightarrow$ the φ_α are chosen to be measurable. We defer general measurability considerations to Section 4.

3. A special group structure. We now discuss a special group structure for which the characterization in Section 2 assumes a particularly interesting form. This structure is common to many equivariant estimation problems. For another reference to the same structure, embedded in a topological context, see [21], Theorem 6.

Suppose there is a further group H of one-one bimeasurable transformations of $\mathfrak{X}$ onto itself such that

(i) Every element of G commutes with every element of H.

(ii) If $g \varepsilon G$, $h \varepsilon H$, $x \varepsilon \mathfrak{X}$ and $gx = hx$, then $gx = x$.

(iii) The group $K = GH$ is transitive on $\mathfrak{X}$.

Property (ii) is a strong requirement that G and H be disjoint. It clearly implies that $G \cap H = \{e\}$, where e is the identity map on $\mathfrak{X}$. Moreover, (i) and (ii) insure that $K = GH$ is the direct product of G and H. Under (iii), $\mathfrak{X}$ is homogeneous regarding K. Thus all of the isotropy subgroups K_x are isomorphic and $\mathfrak{X}$ is itself isomorphic to all of the factor sets K/K_x. (That is, there is a one-one relation between $\mathfrak{X}$ and any K/K_x.) Below, we specify the natural isomorphism; the one that induces on the factor set the natural action of K. This isomorphism is equivariant with respect to K acting on $\mathfrak{X}$ and K/K_x. In addition, (i) and (ii) insure that K/K_x is a direct product as well:

1. PROPOSITION. *Under conditions* (i)–(iii), $K/K_x = G/G_x \cdot H/H_x$, *where* (i) *and* (ii) *insure that multiplication of cosets is unambiguously defined.*

PROOF. By (i), if $k = gh \,\varepsilon\, K_x$, then $g \,\varepsilon\, G_x$ and $h \,\varepsilon\, H_x$. Thus $K_x = G_x H_x$ and if $k' = g'h' \,\varepsilon\, K$, $k'K_x = g'G_x h'H_x$. This representation of K_x-cosets is unambiguous, since $ghK_x = g'h'K_x \Rightarrow g^{-1}g'h^{-1}h' \,\varepsilon\, K_x \Rightarrow gG_x = g'G_x$ and $hH_x = h'H_x$.

We choose a reference point $x_0 \,\varepsilon\, \mathfrak{X}$ and let G_0 etc. denote the isotropy subgroups there. Then the mapping i sending $x = ghx_0$ to gG_0hH_0 is an isomorphism between $\mathfrak{X}$ and $G/G_0 \cdot H/H_0$. This last space is just the direct product of the two factor spaces and we stress this by writing it as $G/G_0 \times H/H_0$. The group K_i induced by i on the product space acts thus: if $k = gh \,\varepsilon\, K$, $k_i(g'G_0, h'H_0) = (gg'G_0, hh'H_0)$. We omit the subscript i when it is not likely to cause confusion. Composing i with projection onto G/G_0 produces a mapping, i_G, of $\mathfrak{X}$ that is (maximally) H-invariant and (G, G_i)-equivariant. (At $x = ghx_0$, $i_G(x) = gG_0$.) A dual statement holds for H and, of course, $i = i_G \times i_H$. (The ranges of the maximal invariants i_G and i_H are chosen for convenience to be the factor sets. Equivalent sets, such as the orbits Gx_0 and Hx_0 can and will be used when preferable.

In view of the isomorphism between $\mathfrak{X}$ and $G/G_0 \times H/H_0$, there is an exact correspondence between equivariant mappings (into Θ) of the two sets, related by i. We will work with the product space and interpret results obtained for the latter in terms of $\mathfrak{X}$. Since (G, G^*)-equivariance is desired and $g \,\varepsilon\, G$ acting on the product space leaves the H-coordinate fixed, the problem reduces to studying (G, G^*)-equivariant mappings of G/G_0, on which G acts transitively. Upon noting that G_0 is the isotropy subgroup of G at $G_0 \,\varepsilon\, G/G_0$, we obtain as a corollary to Theorem 2.1:

2. COROLLARY. *These exist* (G, G^*)-*equivariant mappings* $\varphi: G/G_0 \to \Theta \Leftrightarrow \Theta_0 = \{\theta: (G_0)^* \,\theta = \theta\}$ *is not empty. Moreover,* φ *is determined by its value at any point of* G/G_0, G_0 *say*: $\varphi(gG_0) = g^*\varphi(G_0)$, *where* $\varphi(G_0) \,\varepsilon\, \Theta_0$.

Thus the (G, G^*)-equivariant mappings of $G/G_0 \times H/H_0$ into Θ are obtained by choosing for each $hH_0 \,\varepsilon\, H/H_0$ a (G, G^*)-equivariant mapping $\varphi(\cdot, hH_0)$: $G/G_0 \to \Theta$ and piecing them together. Letting $\varphi_0(\cdot) = \varphi(G_0, \cdot)$; $\varphi_0: H/H_0 \to \Theta_0$ and

$$\varphi(gG_0, hH_0) = g^*\varphi_0(hH_0). \tag{3.1}$$

Thus under conditions (i)–(iii), H/H_0 is a convenient cross-section for the G-orbits of the product space. In $\mathfrak{X}$ itself, this becomes Hx_0 and the characterization becomes: to every (G, G^*)-equivariant mapping $\varphi: \mathfrak{X} \to \Theta$ there corresponds a mapping $\varphi_0: Hx_0 \to \Theta_0$ which uniquely determines φ and at $x = ghx_0$,

$$\varphi x = g^*\varphi_0 hx_0. \tag{3.2}$$

4. Measurable equivariant mappings. We reintroduce into our considerations the σ-fields $\mathcal{A}$ and $\mathcal{B}$ for $\mathcal{X}$ and Θ respectively and seek to characterize the equivariant mappings defined by (3.2) that are measurable. Under the conditions given below, the requirement becomes simply that φ_0 be measurable. It is here that the product space representation becomes particularly useful.

Let $\mathcal{A}_\times = i(\mathcal{A})$ denote the measurable structure for $\mathcal{X}$ transferred to the product space. Let $\mathcal{A}_G$ (resp. $\mathcal{A}_H$) be the σ-field induced on G/G_0 by i_G (resp. on H/H_0 by i_H).

1. THEOREM. *Under the conditions*

(a) $\mathcal{A}_\times = \mathcal{A}_G \times \mathcal{A}_H$ *and*

(b) *the mapping of* $G/G_0 \times \Theta_0$ *into* Θ *defined by* $(gG_0, \theta) \rightarrow g^*\theta$ *is* $\mathcal{A}_G \times \mathcal{B}$ *measurable,*

an equivariant mapping φ, *necessarily of the form* (3.1), *is* $(\mathcal{A}, \mathcal{B})$ *measurable* $\Leftrightarrow \varphi_0$ *is* $(\mathcal{A}_H, \mathcal{B})$ *measurable.*

PROOF. The forward implication follows from (a) since $\varphi_0(\cdot) = \varphi(G_0, \cdot)$. Conversely, choosing $B \varepsilon \mathcal{B}:\{(gG_0, hH_0):\varphi(gG_0, hH_0) \varepsilon B\} = \{(gG_0, hH_0): g^*\varphi_0(hH_0) \varepsilon B\} = (e \times \varphi_0)^{-1}\{(gG_0, \theta) \varepsilon G/G_0 \times \Theta_0: g^*\theta \varepsilon B\}$, where e is the identity map on G/G_0. That the last set in braces belongs to $\mathcal{A}_G \times \mathcal{B}$ is precisely condition (b); the conclusion follows since $e \times \varphi_0$ is $\mathcal{A}_G \times \mathcal{A}_H$ measurable.

In some applications, it happens that G_0 is trivial. Then $\mathcal{A}_G$ is a σ-field for G itself and condition (b) requires that the mapping $(g, \theta) \rightarrow g^*\theta$ be $\mathcal{A}_G \times \mathcal{B}$ measurable. Cf. [14], p. 225, Theorem 4.

We may interpret Theorem 1 in terms of $\mathcal{X}$, thereby removing the necessity for considering a measurable structure on G/G_0. In condition (a), $\mathcal{A}_\times$ is replaced by $\mathcal{A}$ and $\mathcal{A}_G$ and $\mathcal{A}_H$ become induced σ-fields for Gx_0 and Hx_0 respectively. Condition (b) is then that the mapping $(gx_0, \theta) \rightarrow g^*\theta$ be $\mathcal{A}_G \times \mathcal{B}$ measurable. (Note that this mapping is well-defined only if $\theta \varepsilon \Theta_0$.) The theorem then states that under conditions (a) and (b), all mappings of the form (3.2) having measurable φ_0 are measurable and conversely.

5. Examples. Some examples illustrating the preceding ideas are presented. Throughout, $\mathcal{X}$ is essentially R^n, Euclidean n-space and X represents a vector of independent and indentically distributed (IID) random variables. Where convenient, we identify the collection $\mathcal{P}$ of power-product measures with the corresponding family of distributions on the real line. G is correspondingly abused when its elements act coordinatewise in the same manner.

1. Let $\mathcal{P}$ be the family of all continuous distributions on the real line having unique medians. For $P \varepsilon \mathcal{P}$, let $\vartheta(P) = F_P^{-1}(\frac{1}{2})$, where F_P is the CDF associated with P. $\vartheta(P)$ is thus the median associated with P and $\Theta = R$. The group G of real-valued one-one onto increasing functions on R (acting coordinatewise on R^n) leaves $\mathcal{P}$ invariant. It is easily checked that ϑ commutes with $\bar{G}$ and G^* is just G acting on $\Theta = R$: i.e., $\vartheta(\bar{g}P) = F_{\bar{g}P}^{-1}(\frac{1}{2}) = (F_P g^{-1})^{-1}(\frac{1}{2}) = g\vartheta(P)$. We may then require estimators of the median to be (G, G^*)-equivariant.

A group H satisfying conditions (i)–(iii) of Section 3 is the group of permutations on n letters, applied to the coordinates of X. This is easily checked; in order that $K = GH$ be transitive we remove from R^n the $\mathcal{P}$-null set of points with two or more equal coordinates. Choosing $x_0 = (1, 2, \cdots, n)$, if $x = ghx_0$, then $(g(1), \cdots, g(n)) = \langle x \rangle$ is the coordinates of x arranged in increasing order and h rearranges the coordinates to give x. The coset gG_0 is thus identified with $\langle x \rangle$, the common values of its elements on $\{1, \cdots, n\}$ while H has only trivial isotropy subgroups. We may thus choose $i_G(x) = \langle x \rangle$ (i.e., replacing the range G/G_0 by Gx_0) and $i_H(x) = h_x$, where $h_x\langle x \rangle = x$. We may make explicit the representation (3.2) upon delineating Θ_0: G_0 contains all transformations that fix $\{1, \cdots n\}$ pointwise but are otherwise arbitrary. It is clear that θ is fixed by $(G_0)^*$ only if $\theta \in \{1, \cdots, n\}$. Thus $\Theta_0 = \{1, \cdots, n\}$ and from (3.2), $\varphi x = g\varphi_0 hx_0$, where $x = ghx_0$ and $\varphi_0 hx_0 \in \{1, \cdots, n\}$. I.e., if $\varphi_0 hx_0 = m$, $\varphi x = g(m)$, the mth coordinate of $\langle x \rangle$. Thus equivariant estimators of medians (or any other fractile) must be order statistics.

In this characterization, the choice of order statistic is allowed to depend on h_x; i.e., on how x is rearranged to get $\langle x \rangle$. This presumably irrelevant flexibility disappears under a further equivariance requirement: H also leaves $\mathcal{P}$ invariant (pointwise) and we may thus require φ to be (K, K^*)-equivariant. (Note: since $\bar{H}$ is trivial, $\bar{K} = \bar{G} \Rightarrow K^* = G^*$.) Then in addition, φ must satisfy $\varphi hx = \varphi x$ for all h in H, hence φ_0 must be constant. If $\varphi_0 \equiv m$, then the equivariant estimator is the mth order statistic. A result of this nature was obtained by Loynes in [15]. It is essentially shown there that merely requiring $(g^*)^{-1}\varphi gX$ and φX to have the same distribution requires φX to be almost surely an order statistic.

We note that all such estimators are measurable. This simply reflects the fact that the selection of an order statistic (φ_0) is measurable and that there are a finite number of orbits.

2. Let $\mathcal{P}$ be a location and scale family generated by the continuous CDF $F[t]$: $\mathcal{P} = \{F[(t - \mu)/\sigma]: -\infty < \mu < \infty, \sigma > 0\}$. ϑ maps $P \in \mathcal{P}$ into the corresponding value of (μ, σ); hence $\Theta \subset R^2$. Let G be the group of location and scale transformations: $g = (a, b)$ sends $x \to a + bx$ (coordinatewise); $-\infty < a < \infty$, $b > 0$. $\bar{G}$ is isomorphic to G and ϑ, being one-one automatically commutes with $\bar{G}$. G^* is thus Θ acting on itself: $(a, b)^*(\mu, \sigma) = (a + b\mu, b\sigma)$. Upon eliminating the $\mathcal{P}$-null set of points having all coordinates equal, the group H of $n \times n$ orthogonal matrices that fix $\mathbf{1} = (1, \cdots, 1) \in R^n$ satisfies conditions (i)–(iii): (i) is easily checked; to verify (ii), note that every $h \in H$ preserves the lengths of projections along and orthogonal to $\mathbf{1}$. In order that $g \in G$ do likewise, we must have $b = 1$ and $a = 0$. To verify (iii), note that starting with an arbitrary x, its component orthogonal to $\mathbf{1}$ may be modified in an arbitrary manner by a combination of a rotation in H and a scale change; the resulting component along $\mathbf{1}$ may then be changed arbitrarily by shifting.

The isotropy subgroups of G are all trivial. We choose x_0 to be of unit length and orthogonal to $\mathbf{1}$. If $x = ghx_0$ with $g = (a, b)$, then $x = a\mathbf{1} + bhx_0$ where hx_0 is orthogonal to $\mathbf{1}$. Thus $a = \bar{x}$ is the length of the component of x along $\mathbf{1}$, hx_0 is the direction of the complementary component and $b = s(x)$ is its length ($\bar{x} = (\sum x_j)/n$ and $s(x) = (\sum (x_j - \bar{x})^2)^{\frac{1}{2}}$.) It is convenient here to choose the range of i_H to be Hx_0. Thus at $x = ghx_0$, $i_G(x) = g = (\bar{x}, s(x))$ and $i_H(x) = hx_0 = ((x_1 - \bar{x})/s(x), \cdots, (x_n - \bar{x})/s(x))$. Since G_0 is trivial, $\Theta_0 = \Theta \subset R^2$ and (3.2) shows that equivariant estimators are of the form $\varphi x = (\bar{x}, s(x))^* \varphi_0(i_H(x))$, where φ_0 assumes values in Θ. Denoting its components by φ_μ and φ_σ, we see that $\varphi(x) = [\bar{x} + s(x)\varphi_\mu(i_H(x)), s(x)\varphi_\sigma(i_H(x))]$, where φ_μ and φ_σ are essentially arbitrary mappings of the unit sphere in R^{n-1} into R and $(0, \infty)$ respectively. Pitman's equivariant estimators of location and scale are, with a bit of rearranging, seen to be of this form [17].

Introducing measurability requirements simply constrains φ_μ and φ_σ to be measurable. This is easly seen directly, but it also follows upon verifying that conditions (a) and (b) of Theorem 4.1 hold. We note that with the present choice of ranges for i_G and i_H, $\mathcal{A}_G$ is the Borel subsets of G and $\mathcal{A}_H$, those of the unit sphere in R^{n-1}. Also, as G_0 is trivial, condition (b) is fulfilled since the map sending $((a, b), (\mu, \sigma))$ to $(a + b\mu, b\sigma)$ is jointly measurable.

3. We continue example 2 by supposing that F is the unit normal CDF. Then X is a vector of IID normal variables and H also leaves $\mathcal{P}$ invariant (pointwise). ϑ commutes with $\bar{K} = \bar{G}$, $K^* = G^*$ and we may require (K, G^*)-equivariance. The added requirement, that $\varphi hx = \varphi x$ for all h in H reduces φ_μ and φ_σ to constants. In fact, we may augment K by including the transformation $X \to -X$ which leaves $\mathcal{P}$ invariant; then $\varphi_\mu \equiv 0$. Thus $\varphi x = (\bar{x}, cs(x))$, where $c \equiv \varphi_\sigma$ is a positive constant. Hence one obtains estimators that are (natural) functions of the sufficient statistic, although no sufficiency considerations were explicitly introduced. This becomes less surprising in the light of [1] where Basu shows that certain sufficient statistics are maximal invariants.

We discuss now an example where a direct appeal to the construction of Section 2 seems necessary.

4. Let $\mathcal{P}$ be the family of all distributions on the real line. For $P \varepsilon \mathcal{P}$, let $\vartheta(P) = F_P$, the associated CDF. Θ is the set of CDF's on R. The groups G and H of example 1 leave $\mathcal{P}$ invariant and as ϑ is one-one, it automatically commutes with $\bar{K} = \bar{G}$: if $k = gh$, $k^* = g^*$ sends $F \varepsilon \Theta$ into $g^*F = Fg^{-1}$. Thus we may require an estimator of the CDF to be (K, G^*)-equivariant. We cannot, as in example 1, remove a $\mathcal{P}$-null set to make K transitive. Nor does there seem to be a further H-group which would produce transitivity. As there are but a finite number of K-orbits, we may use construction of Section 2 and not worry about measurability. The different K-orbits correspond to the possible patterns of repetitions of coordinates. Thus $(1, 2, \cdots, n)$, $(1, 1, 2, \cdots, n - 1)$ and $(1, 1, 2, 2, 3, \cdots, n - 2)$ are in different K-orbits. We may choose as a reference point in each orbit a generic point such as $(1, \cdots, 2, \cdots, \cdots, j) = x_\alpha$, where the multiplicities of the indices $\{1, 2, \cdots, j\}$ characterize the orbit. The nature of Θ_α follows from that of G_α: $g \varepsilon G_\alpha$ if $gx_\alpha = x_\alpha$; i.e., if g fixes $\{1, \cdots, j\}$. If $F \varepsilon \Theta_\alpha$, then for all g in G_α, $Fg^{-1} = F$. This cannot happen if F increases on $R - \{1, \cdots, j\}$; hence Θ_α contains precisely those CDF's that increase only on $\{1, \cdots, j\}$. Upon choosing $F_\alpha \varepsilon \Theta_\alpha$ for each α, one sees that at $x = ghx_\alpha$, $\varphi x = F_\alpha g^{-1}$. Since $gx_\alpha = \langle x \rangle$, $F_\alpha g^{-1}$ increases on the coordinates of $\langle x \rangle$. Thus an

equivariant estimate of the CDF is one that increases only on the order statistics. Note that the size of the jumps can (indeed, must) depend on the orbit of x, i.e., the number of different coordinates x has. This result is essentially obtained in [14], pp. 246–247 by *ad hoc* methods. Note that G may be considerably diminished without altering the restricted nature of equivariant estimators of the CDF.

It is tempting to try to make K transitive in this example by defining $\mathcal{P}$ to contain only continuous distributions. Then, as in example 1, all but one K-orbit is $\mathcal{P}$-null. However, Θ must then be properly taken to be the continuous CDF's (i.e., the range of ϑ, only there can G^* be defined) and Θ_0 would be empty. That is, there are no equivariant estimates of continuous CDF's that are themselves of this nature.

This phenomenon is exhibited in another natural estimation situation in which no equivariant estimator exists.

5. Let $\mathcal{P}$ be the set of all absolutely continuous distributions on R and let $\vartheta(P) = dP/d\lambda = f_P$, where λ denotes Lebesgue measure. Thus $\Theta \subset L_1(\lambda)$. Let G be the group of all diffeomorphisms of R, the one-one monotone functions with non-vanishing derivative. (G acts coordinatewise). ϑ, being one-one, commutes with $\bar{G}$ and if $g \varepsilon G, g^*f = (fg^{-1})|dg^{-1}/d\lambda|$. As in examples 1 and 4, H may be taken to be the permutations. We choose $x_0 = (1, \cdots, n)$ and consider Θ_0. G_0 is again the subgroup fixing $\{1, \cdots, n\}$ pointwise and if every $g^* \varepsilon (G_0)^*$ fixes $f \varepsilon \Theta$, they fix the corresponding CDF in that $Fg^{-1} = F$. Again, this means that F can only increase on $\{1, \cdots, n\}$ and hence is not absolutely continuous. Thus Θ_0 is empty and equivariant estimation of the density is not possible. (One may, in a sense, beg the issue by adopting a more general view of "density", allowing Schwarz-distributions as derivatives of discrete CDF's. This does not seem compelling from a statistical point of view.)

6. Discussion. The group structure presented in Section 3 has desirable consequences in various statistical applications. We mention first the problem of constructing a function on $\mathfrak{X}$ that is maximally G-invariant. In principle, such always exist, since $\mathfrak{X}$ can always be mapped into the set of G-orbits or any cross-section of them. Under conditions (i)–(iii), a natural cross-section is an H-orbit, Hx_0 say. Thus $i_H(x)$ (taking its range to be Hx_0) provides a specific maximal G-invariant. See, for example, the representation of i_H in example 2 above. As $i_H(x) = hHx_0$ at $x = ghx_0$, $i_H(h'x) = h'i_H(x)$; i.e., a (maximal) G-invariant commutes with H and vice versa. This property is used inherently in the considerations of Section 3.

This last fact sometimes underlies a phenomenon observed in situations involving a sufficiency reduction where invariance considerations are present: often the structure $(X, \mathcal{P}, G_X)$ can be reduced to $(T(X), \mathcal{P}_T, G_T)$, where $T(X)$ is sufficient for $\mathcal{P}$, $\mathcal{P}_T$ is the induced set of distributions for T and G_T is a group of transformations induced on T. See examples 1 and 3 above, as well as [14], Chapter 6, (examples 5 and 6, e.g.). Given the sufficiency of $T(X)$, the part of the reduction that is not automatic is the passage from G_X to G_T. When possible, it often is explained by the fact that there are sugbroups G and H satisfying (i)–(iii) with $G_X \subset G$ and $T(X) = i_G(X)$, the maximal H-invariant. T then commutes with G and therefore with any subgroup. There is an extensive discussion based on this phenomenon in [9]. In [7], Fraser considers a special group structure which provides a natural fiducial distribution for the parameter ϑ; more specifically, a natural pivotal quantity. Fraser requires that Θ, $\mathcal{P}$ and G be isomorphic (i.e., that $\bar{G}$ be isomorphic to G and transitive on $\mathcal{P}$ and that ϑ be one-one). His fiducial argument generally involves conditioning on an ancillary statistic which, Fraser points out, in this situation may be thought of as the G-orbit of X. Under the structure of Section 3, one requires explicitly the conditional distribution of $i_G(X)$ given $i_H(X)$. Stone [19] also requires a representation of the form $X = (a(X), b(X))$, where the range of a is isomorphic to G and $b(X)$ is ancillary. i_G and i_H provide such a pair of functions when the isotropy subgroups of G are trivial.

Finally, a word about randomized estimators. One may adopt the view that a randomized estimator merely depends on augmented data (X, U), where U is an extraneous random variable distributed independently of $\mathfrak{X}$ on a sample space $\mathfrak{U}$. We may extend G to $\mathfrak{X} \times \mathfrak{U}$ by defining $g(x, u) = (gx, u)$ for all g, x and u. The equivariance requirement is then $\varphi(gx, u) = g^*\varphi(x, u)$; hence for fixed u, φ is equivariant in x and the preceding applies. One may then piece together equivariant $\varphi(\cdot, u)$, one for each u in $\mathfrak{U}$. Of course it is desirable that the resulting $\varphi(\cdot, \cdot)$ be jointly measurable. Under conditions (a) and (b) of Section 4, this may be accomplished by choosing $\varphi_0: Hx_0 \times \mathfrak{U} \to \Theta_0$ to be jointly measurable and defining $\varphi(x, u) = g^*\varphi_0(hHx_0, u)$ at $x = ghx_0$. Thus one first randomizes over Θ_0 (with φ_0) and then proceeds equivariantly.

7. Acknowledgment. The author is indebted to R. A. Wijsman for many constructive comments on this work. In particular, he suggested the measurability characterizations presented in Sections 2 and 4.

REFERENCES

[1] Basu, D. (1965). On sufficiency and invariance. Unpublished.
[2] Blackwell, D. and Girshick, M. A. (1954). *Theory of Games and Statistical Decisions.* Wiley, New York.
[3] Blyth, C. (1951). Admissibility of minimax procedures. *Ann. Math. Statist.* **22** 22–42.
[4] Brown, L. D. (1966). On the admissibility of invariant estimators of one or more location parameters. *Ann. Math. Statist.* **37** 1087–1136.
[5] Chen, H. J. (1964). On minimax invariant estimates of scale and location parameters. *Sci. Sinica* **13** 1569–1586.
[6] Chen, P. (1964). Minimax estimates of parameters of distributions belonging to the exponential family. *Chinese Math.* **5** 277–299.
[7] Fraser, D. A. S. (1961). The fiducial method and invariance. *Biometrika* **48** 261–280.
[8] Girshick, M. A. and Savage, L. J. (1951). Bayes and minimax estimates for quadratic loss functions. *Proc. Second Berkeley Symp. Math. Statist. Prob.* 53–74. Univ. of California Press.
[9] 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.

[10] JAMES W. and STEIN C. (1961). Estimation with quadratic loss. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* **1** 361–379. Univ. of California Press.

[11] KARLIN, S. (1953). The theory of infinite games. *Ann. Math.* **58** 371–401.

[12] KIEFER, J. (1957). Invariance, minimax sequential estimation and continuous time processes. *Ann. Math. Statist.* **28** 573–601.

[13] KUDŌ, H. (1955). On minimax invariant estimates of the transformation parameter. *Nat. Sci. Rept. Ochanomizu Univ., Tokyo* **6** 31–73.

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

[15] LOYNES, R. M. (1966). Estimation of quantiles. *J. Roy. Statist. Soc. Ser. B* **28** 497–512.

[16] MONTGOMERY, D. and ZIPPIN, L. (1955). *Topological Transformation Groups.* Interscience, New York.

[17] PITMAN, E. J. G. (1939). The estimation of location and scale parameters of a continuous population of any given form. *Biometrika* **31** 391–421.

[18] STEIN, C. (1959). The admissibility of Pitman's estimator of a single location parameter. *Ann. Math. Statist.* **30** 970–979.

[19] STONE, M. (1965). Right Haar measure for convergence in probability to quasi posterior distributions. *Ann. Math. Statist.* **36** 440–453.

[20] WESLER, O. (1959). Invariance theory and a modified minimax principle. *Ann. Math. Statist.* **30** 1–20.

[21] WIJSMAN, R. A. (1967). Cross-sections of orbits and their application to densities of maximal invariants. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* Univ. of California Press.