The Annals of Mathematical Statistics
1970, Vol. 41, No. 5, 1708–1716

THE APPLICATION OF INVARIANCE TO UNBIASED ESTIMATION[1]

BY MORRIS L. EATON AND CARL N. MORRIS

University of Chicago and The RAND Corporation

1. Introduction and Summary. A number of articles concerned with the problem of finding minimum variance unbiased estimates (MVUE) of certain "non-standard" parametric functions have appeared in the statistical literature. When a complete sufficient statistic (c.s.s.) exists, the Rao–Blackwell Theorem gives a method of constructing the MVUE of an estimable parametric function based on any unbiased estimate. A parametric function $\varphi(\theta)$ is called *estimable* if there exists an unbiased estimate.

As an example, let $X_1, \cdots, X_n$ be independent and identically distributed (i.i.d.), X_i being distributed according to a p-dimensional multivariate normal distribution with mean $\mu \in R^p$ ($R^p = p$-dimensional Euclidean space) and covariance matrix $\Sigma : p \times p$. This is denoted by $\mathscr{L}(X_i) = \mathscr{N}_p(\mu, \Sigma)$, $\mathscr{L}(X)$ reading "the law of X." All vectors in this paper are column vectors. For $p = 1$, Kolmogorov (1950) solved the problem of finding the MVUE for the parametric function

$$\varphi_t(\mu, \sigma^2) = P_{\mu,\sigma^2}(X_1 \geqq t),$$

μ and σ^2 both unknown. Kolmogorov's method, which was to calculate the conditional distribution of X_1 given the c.s.s. and apply the Rao–Blackwell Theorem to the indicator function $I_{[t,\infty)}(X_1)$, was sufficiently general to be applicable to other estimable parametric functions for this family of distributions. Later Lieberman and Resnikoff (1955) gave an independent solution to this problem using the same method. Barton (1961) and A. P. Basu (1964) solved this and other problems using the same approach.

An alternative method for finding MVUE's when the distribution of the c.s.s. is given, is the transform (LaPlace, Mellin, etc.) method. When applicable, this method does not require having an initial unbiased estimate. The transform method was used by Tate (1959) for distributions involving location and scale parameters and by Washio, Morimoto and Ikeda (1956) for the one-parameter Koopman–Darmois family. Olkin and Pratt (1958) also determined MVUE's of certain correlation coefficients using the LaPlace Transform.

Neyman and Scott (1960) developed a third method, which they term "the expansion method," for producing the MVUE of certain parametric functions. Their applications were restricted to univariate normal distributions.

An obvious alternative to the above methods is to exhibit a function of the c.s.s. and verify that it is unbiased. This method was used by Ghurye and Olkin (1969) to estimate density functions of the multivariate normal and the Wishart distribu-

Received November 24, 1969.

[1] This research was supported by the United States Air Force under Project RAND.

tion with various parametric assumptions. As it is not constructive, this method may be hard to apply.

E. Lehmann (unpublished lecture notes prior to 1964) and Sathe and Varde (1969) used a theorem due to D. Basu (1955) to greatly simplify the application of the Rao–Blackwell Theorem. Our Theorem 2.1, which permits other (including multivariate) applications, is a generalization of their approach. Theorem 2.1 is applicable when a given unbiased estimate may be written as a function of the c.s.s. and an ancillary statistic. An ancillary statistic is one whose distribution does not depend on the parameter.

The preponderance of applications of all the above references have been to families of distributions which are invariant under a group of transformations. These applications are in the domain of our Theorem 2.2 which gives conditions under which a group of transformations may be used to construct the data as a function of a c.s.s. and an ancillary statistic determined by the group structure and the c.s.s. In most cases, this ancillary statistic will be a maximal invariant statistic. The group structure essentially provides an easy method for representing the conditional distribution of the data given the c.s.s. in terms of the marginal distribution of the ancillary statistic. This greatly simplifies the application of the Rao–Blackwell Theorem to a given unbiased estimate.

Section 3 illustrates the application of Theorem 2.2. In Example 1, the $\mathcal{N}_p(\mu, \Sigma)$ distribution is considered, μ and Σ both unknown. The distribution of a maximal invariant is characterized and used to represent in integral form the MVUE of any estimable parametric function. Application is given to set probabilities and to estimating the $\mathcal{N}_p(\mu, \Sigma)$ density. Example 2 is concerned with an application to U-statistics.

2. The Main Theorem. Let X be a random variable taking values in an abstract sample space $(\mathscr{X}, \mathscr{B})$ and suppose the distribution of X belongs to a specified family $\mathscr{P} = \{P_\theta \mid \theta \in \Theta\}$ of probability distributions on $(\mathscr{X}, \mathscr{B})$. A real-valued function $f(X)$ is given. The problem is to determine the MVUE of $\varphi(\theta)$ defined by

$$\varphi(\theta) \equiv E_\theta f(X). \tag{2.1}$$

Assume that $T \equiv T(X)$ is a c.s.s. for $\mathscr{P}$. The following result is due to D. Basu (1955).

THEOREM (Basu). *Suppose Y, a function on $\mathscr{X}$ to a space $\mathscr{Y}$, is an ancillary statistic. Then $T(X)$ and $Y(X)$ are stochastically independent for each $\theta \in \Theta$.*

THEOREM 2.1. *Suppose there exists a function Z on $(\mathscr{X}, \mathscr{B})$ such that*

(i) *$Z(X)$ is ancillary, and*
(ii) *there exists a function $W = W(Z, T)$ such that $f(X) = W(Z(X), T(X))$.*

Then the MVUE *for $\varphi(\theta)$ is*

$$f^*(T) \equiv E_Z W(Z, T) \tag{2.2}$$

where E_Z denotes expectation with respect to the marginal distribution of Z.

PROOF. From the Rao–Blackwell Theorem,

$$f^*(t) = E(f(X) \mid T(X) = t) \tag{2.3}$$

is the MVUE for $\varphi(\theta)$. Therefore,

$$\begin{aligned} f^*(t) &= E(W(Z, T(X)) \mid T(X) = t) \\ &= E(W(Z, t) \mid T(X) = t) \\ &= E_Z W(Z, t) \end{aligned} \tag{2.4}$$

since Z and T are independent by Basu's Theorem. This completes the proof.

The main problem in applying Theorem 2.1 is that of finding the functions Z and W. In many applications, one also has a group of transformations G which acts on the sample space $\mathscr{X}$ and preserves the family $\mathscr{P}$. That is, each $g \in G$ is a $1-1$ bimeasurable function on $(\mathscr{X}, \mathscr{B})$ onto $\mathscr{X}$, with the group operation $\circ$ satisfying $(g_1 \circ g_2)(x) = g_1(g_2(x))$. G preserves the family $\mathscr{P}$ means that for each $g \in G$, the probability measure $g \circ P_\theta$, defined by $(g \circ P_\theta)(B) \equiv P_\theta(g^{-1}(B))$, is in $\mathscr{P}$. G then induces a group $\bar{G}$ of transformations on Θ in a natural way, $\bar{g}\theta$ being defined by the relation $g \circ P_\theta = P_{\bar{g}\theta}$. For further discussion, the reader is referred to Lehmann (1959).

The next theorem uses the group structure to determine the functions Z and W of Theorem 2.1. As above, let X take values in $(\mathscr{X}, \mathscr{B})$ with a distribution in $\mathscr{P} = \{P_\theta \mid \theta \in \Theta\}$ and suppose the group G acts on $\mathscr{X}$ and preserves $\mathscr{P}$. Assume that $T: \mathscr{X} \to G$ is a c.s.s. which takes values in G. The value in G of T at X will be written T_X. The notation T_X is used, instead of $T(X)$, because T_X is an element of G and is thus a function on $\mathscr{X}$ to $\mathscr{X}$ in addition to being a c.s.s. If $y \in \mathscr{X}$, the value of $T_X \in G$ is then denoted by $T_X(y)$.

THEOREM 2.2. *Let $Z(X) \equiv T_X^{-1}(X)$ where T_X^{-1} is the group inverse of $T_X \in G$. If $Z(X)$ is an ancillary statistic, then the* MVUE *for $\varphi(\theta)$ is given by*

$$\begin{aligned} f^*(T_X) &= E_Z f(T_X(Z)) \\ &= \int f(T_X(z))\zeta(dz) \end{aligned} \tag{2.5}$$

where ζ is the marginal distribution of Z.

PROOF. $f(X) = f(T_X(Z))$ and therefore (ii) of Theorem 2.1 is satisfied. Since Z is ancillary and T_X is a c.s.s., the conclusion follows.

REMARK 1. A sufficient condition that $Z(X)$ of Theorem 2.2 be ancillary is that $\bar{G}$ be transitive on Θ and $Z(X)$ be invariant (Lehmann (1959)). $\bar{G}$ is transitive on Θ means for each $\theta_1, \theta_2 \in \Theta$, there exists $\bar{g} \in \bar{G}$ such that $\theta_2 = \bar{g}(\theta_1)$. $Z(X)$ is invariant means $Z(g(x)) = Z(x)$ for all $g \in G$, $x \in \mathscr{X}$. The condition

$$T_{g(x)} = g \circ T_x \tag{2.6}$$

for all $g \in G$, $x \in \mathscr{X}$ implies that $Z(X)$ is invariant, since

$$\begin{aligned} Z(g(x)) \equiv T_{g(x)}^{-1}(g(x)) &= (g \circ T_x)^{-1}(g(x)) \\ &= T_x^{-1} \circ g^{-1} \circ g(x) = Z(x). \end{aligned}$$

In many applications, one can choose T so that (2.6) holds.

REMARK 2. Theorem 2.2 essentially gives a constructive method for determining a version of the conditional distribution of the data given the c.s.s., and then straightforwardly applies the Rao–Blackwell Theorem. Formally, this conditional probability is given for $B \in \mathscr{B}$ by

$$\begin{aligned} P(X \in B \mid T_X = t) &= P(T_X T_X{}^{-1}(X) \in B \mid T_X = t) \\ &= P(t(Z) \in B \mid T_X = t) \\ &= P(Z \in t^{-1}(B) \mid T_X = t) \\ &= P(Z \in t^{-1}(B)) \qquad \text{since } Z \text{ and } T_X \text{ are independent} \\ &= \zeta(t^{-1}(B)). \end{aligned}$$

Also, Theorem 2.2 shows that if $Z(X)$ is ancillary, then $E(F(X) \mid T(X) = t) = E_Z F(t(Z))$ where F is any integrable function of X. This result can be used to "Blackwell–Rao" any estimator and for convex loss functions, this gives an improved estimator.

If $Z(X)$ is invariant, then the assumptions of Theorem 2.2 imply that $Z(X)$ is a maximal invariant in $\mathscr{X}$ under G. (See Lehmann (1959) for the definition of a maximal invariant.) In this case, the distribution of Z can sometimes be obtained by averaging the distribution of X over G with respect to Haar measure on G (Schwartz (1966) and Wijsman (1965)).

One can also apply the arguments of this section to the problem of unbiasedly estimating a vector-valued parametric function. Theorems 2.1 and 2.2 are simply applied coordinatewise, and the only unbiased estimate based on a c.s.s. is still given formally by (2.2) or (2.5).

In many cases the groups G and $\bar{G}$ will be isomorphic rather than simply homomorphic. The following example illustrates that the choice of G may be important in verifying (2.6).

Define $e = (1, \cdots, 1)'$ as the n-vector of units and let $\mathscr{L}(X) = \mathscr{N}_n(\theta e, I)$ with $\theta \in \Theta \equiv R$. It is required to find the MVUE of $\varphi(\theta) = E_\theta f(X)$. This family of distributions is invariant under the additive group $G \equiv \{g_a \colon R^n \to R^n \text{ by } g_a(x) = x + ae\}$, the group operation being $g_a \circ g_b = g_{a+b}$. A c.s.s. $T \colon R^n \to G$ is given by $T_x = g_{\bar{x}}$, $\bar{x} \equiv \sum x_i/n$. Clearly $T_{g_a}(x) = g_{\bar{x}+a} = g_a \circ T_x$ so that (2.6) holds.

$\bar{G} \equiv \{\bar{g}_a \colon \bar{g}_a(\theta) = \theta + a\}$ is induced by G on Θ and is isomorphic to G. Since $\bar{G}$ acts transitively on Θ, $Z(X) = T_X{}^{-1}(X) = X - \bar{X}e$ is ancillary and Theorem 2.2 applies. From (2.5), the desired estimate is $\int f(\bar{x}e + z)\, d\zeta(z)$, ζ being the distribution of Z which is singular, $\mathscr{N}_n(0, I - ee'/n)$.

Define $\mathscr{O}(n)$ to be the group of $n \times n$ orthogonal matrices and $\mathscr{O}_e(n) = \{\Gamma \in \mathscr{O}(n) \colon \Gamma e = e\}$. The family of distributions of X is also invariant under the group

$$G^* = \{g^*_{\Gamma,a} \colon \Gamma \in \mathscr{O}_e(n),\ a \in R,\ g^*_{\Gamma,a}(x) \equiv \Gamma x + ae\},$$

with group operation

$$g^*_{\Gamma_2,a_2} \circ g^*_{\Gamma_1,a_1} = g^*_{\Gamma_2\Gamma_1,a_2+a_1}.$$

$\bar{G}^*$ induced on Θ by G^* is $\bar{G}$ and hence G^* and $\bar{G}^*$ are not isomorphic. The statistic $T_X{}^* = (I, \bar{X}e)$ is again a c.s.s. and $T_X{}^{*-1}(X)$ is ancillary, so Theorem 2.2 is applicable. However, (2.6) does not hold for $T_X{}^*$ and the group G^* since $T^*_{g^*(x)}(y) = y+(\bar{x}+a)e$ and $g^* \circ T_x{}^*(y) = \Gamma y + (\bar{x}+a)e$.

3. Examples. In this section, two examples are given which illustrate the use of Theorem 2.2.

EXAMPLE 1. The following example typifies the application of Theorem 2.2. Let $X = (X_1, \cdots, X_n)$ be a $p \times n(n \geqq p+1)$ random matrix whose columns X_i are i.i.d., $\mathscr{L}(X_i) = \mathscr{N}_p(\mu, \Sigma)$, $\mu \in R^p$, $\Sigma > 0$ both unknown.

Denote $G^+_{T,p}$ as the group of $p \times p$ lower triangular matrices with positive diagonal elements. Consider the group $G = \{(A, b): A \in G^+_{T,p}, b \in R^p\}$ with the group operation $(A_2, b_2) \circ (A_1, b_1) \equiv (A_2 A_1, A_2 b_1 + b_2)$. The action of the group G on $\mathscr{X}$ is given by $(A, b)(X) = AX + be'$, e being n-dimensional.

A positive definite matrix S has many square roots B satisfying $S = BB'$. Define the function sqrt (S) to be the unique square root of S in $G^+_{T,p}$. Then sqrt $(ASA') = A$ sqrt (S) if $A \in G^+_{T,p}$. Some square roots do not have this property.

Let $\bar{X} = n^{-1}Xe$ and $S = (X - \bar{X}e')(X - \bar{X}e')'$ so that $(S, \bar{X})$ is a c.s.s. for the family of distributions of X. Then

$$T_X \equiv (B, \bar{X}) \in G \tag{3.1}$$

is a c.s.s. where $B \equiv \text{sqrt}(S)$. A discussion and further references concerning the matrix B are given in Wijsman (1959).

$$T_X{}^{-1} = (B^{-1}, -B^{-1}\bar{X}), \quad \text{so that} \tag{3.2}$$

$$Z(X) \equiv T_X{}^{-1}(X) = B^{-1}(X - \bar{X}e). \tag{3.3}$$

If X is transformed to AX for $A \in G^+_{T,p}$, then $B = \text{sqrt}(S)$ is transformed to AB. Hence, T_X satisfies (2.6). $\bar{G}(= G)$, with group operation on the parameter space given by $(A, b)(\mu, \Sigma) = (A\mu + b, A\Sigma A')$, acts transitively on the parameter space. Thus $Z(X)$ is ancillary and by Theorem 2.2, the MVUE for $Ef(X)$ is

$$f^*(T_X) = f^*(B, \bar{X}) = \int f(Bz + \bar{X}e')\, d\zeta(z) \tag{3.4}$$

where ζ is the distribution of Z.

We now consider the distribution of Z given in (3.3). Setting $P_e = n^{-1}ee'$, $X - \bar{X}e' = X(I - P_e)$ and $S = BB' = X(I-P_e)X'$. Let $\tilde{\Gamma} \in \mathscr{O}(n)$ be represented as $\tilde{\Gamma} = (\Gamma, e/n^{\frac{1}{2}})$ where $\Gamma: n \times (n-1)$ satisfies $\Gamma\Gamma' = I_n - P_e$, $\Gamma'\Gamma = I_{n-1}$ and $\Gamma'e = 0$. Then

$$\tilde{\Gamma}'(I - P_e)\tilde{\Gamma} = \begin{pmatrix} I_{n-1} & 0 \\ 0' & 0 \end{pmatrix}. \tag{3.5}$$

Setting $\tilde{Y} = X\tilde{\Gamma}: p \times n$ and letting $Y: p \times (n-1)$ be the first $n-1$ columns of $\tilde{Y}$ gives

$$BB' = S = YY' \quad \text{and} \tag{3.6}$$

$$X(I - P_e) = (Y, 0)\tilde{\Gamma}' = Y\Gamma'. \tag{3.7}$$

Thus

$$Z = B^{-1}Y\Gamma' \tag{3.8}$$

where $Y = (Y_1, \cdots, Y_{n-1})$ is a $p \times (n-1)$ random matrix whose columns are i.i.d., and $\mathscr{L}(Y_i) = \mathscr{N}_p(0, \Sigma)$.

Now, let $V(Y) \equiv B^{-1}Y$ with $B = \text{sqrt}(YY')$, and note that $V(AY\Delta) = V(Y)\Delta$ if $A \in G_{T,p}^{+}$ and $\Delta \in \mathscr{O}(n-1)$. Since $G_{T,p}^{+}$ acts transitively on the space of covariance matrices, the distribution of V does not depend on Σ. The relation $\mathscr{L}(Y\Delta) = \mathscr{L}(Y)$ for all $\Delta \in \mathscr{O}(n-1)$, implies

$$\mathscr{L}(V) = \mathscr{L}(V\Delta) \tag{3.9}$$

for all $\Delta \in \mathscr{O}(n-1)$. Define $\mathscr{F}(p, n-1)$ as the space of row-orthogonal matrices $M: p \times (n-1)$ satisfying $MM' = I_p$. Then $V = V(Y) \in \mathscr{F}(p, n-1)$. The group $\mathscr{O}(n-1)$ acts transitively on $\mathscr{F}(p, n-1)$ under the action $M \to M\Delta$, $\Delta \in \mathscr{O}(n-1)$, so the distribution of V is uniquely characterized by (3.9) (Nachbin (1965)). This distribution is the "uniform distribution" on $\mathscr{F}(p, n-1)$ under the action of $\mathscr{O}(n-1)$.

Define $\mathscr{F}_e(p, n) = \{M \in \mathscr{F}(p, n): Me = 0\}$. $\mathscr{O}_e(n)$ acts transitively on $\mathscr{F}_e(p, n)$ under the action $M \to M\psi$. From (3.8), $Z \in \mathscr{F}_e(p, n)$ and if $\psi \in \mathscr{O}_e(n)$, then

$$\begin{aligned} Z\psi &= V\Gamma'\psi = V\Gamma'\psi(I-P_e) \\ &= V\Gamma'\psi\Gamma\Gamma' \equiv V\Delta\Gamma' \end{aligned}$$

where $\Delta \equiv \Gamma'\psi\Gamma \in \mathscr{O}(n-1)$. Therefore,

$$\begin{aligned} \mathscr{L}(Z\psi) &= \mathscr{L}(V\Delta\Gamma') \\ &= \mathscr{L}(V\Gamma') \qquad \text{from (3.9)} \\ &= \mathscr{L}(Z) \end{aligned}$$

so that Z has the "uniform distribution" over $\mathscr{F}_e(p, n)$ under the action of $\mathscr{O}_e(n)$.

The preliminary result showing V is uniform on $\mathscr{F}(p, n-1)$ under the action of $\mathscr{O}(n-1)$ is useful for application to problems where μ is known.

In many cases of interest, the function $f(X)$ depends on X_1 only, i.e., $f(X) = h(X_1)$. It is then possible to give a more explicit expression for f^* defined in (3.4). Using the relation $X_1 = X\varepsilon_1$, $\varepsilon_1' \equiv (1, 0, \cdots, 0): 1 \times n$, then

$$\begin{aligned} f(Bz+\bar{X}e') &= h[(Bz+\bar{X}e')\varepsilon_1] \\ &= h(Bz_1+\bar{X}) \end{aligned}$$

with z_1 the first column of z. The distribution of Z_1 is therefore needed. But, $Z_1 = Z\varepsilon_1 = V\Gamma'\varepsilon_1 \equiv V\xi_1$, ξ_1 being the first column of Γ' so that $||\xi_1||^2 = 1-n^{-1}$. Thus,

$$\begin{aligned} \mathscr{L}(Z_1) &= \mathscr{L}(V\xi_1) = \mathscr{L}(V\Delta\xi_1) \\ &= \mathscr{L}([(n-1)/n]^{\frac{1}{2}}V_1) \end{aligned} \tag{3.10}$$

by choosing $\Delta \in \mathcal{O}(n-1)$ so that $\Delta\xi_1 = [(n-1)/n]^{\frac{1}{2}}\varepsilon_1 \in R^{n-1}$. $V_1 \in R^p$ is the first column of V. From the derivation of the distribution of V, V_1 has the same distribution as the first p components of a random vector which is distributed uniformly on $\{x \mid x \in R^{n-1}, \ \|x\| = 1\}$. Hence, $\mathscr{L}(V_1) = \mathscr{L}((U_1, \cdots, U_p)'/\|U\|)$ where $\mathscr{L}(U) = \mathscr{N}_{n-1}(0, I_{n-1})$, and $U' = (U_1, \cdots, U_{n-1})$. Consequently, $\|V_1\|^2$ has a beta distribution with parameters $\frac{1}{2}p$ and $\frac{1}{2}(n-p-1)$. Using this and the relation $\mathscr{L}(V_1) = \mathscr{L}(\Delta V_1)$ for all $\Delta \in \mathcal{O}(p)$, the density of V_1 is given by

$$k(v_1) = c(1-v_1'v_1)^{\frac{1}{2}(n-p-3)} \quad \text{for} \quad v_1'v_1 \leqq 1; \qquad = 0 \quad \text{otherwise} \tag{3.11}$$

where $c = \Gamma(\frac{1}{2}(n-1))[\pi^{\frac{1}{2}p}\Gamma(\frac{1}{2}(n-p-1))]^{-1}$. This expression for $k(v_1)$ is easily verified by using a transformation to polar coordinates (see Tamhankar (1967)).

From (3.10) and (3.11), Z_1 has density function

$$k^*(z_1) = c^*\left(1 - \frac{n}{n-1}z_1'z_1\right)^{\frac{1}{2}(n-p-3)} I_n(z_1) \tag{3.12}$$

where

$$c^* = \left(\frac{n}{n-1}\right)^{\frac{1}{2}p} c,$$

and

$$I_n(z_1) = 1 \quad \text{if} \quad z_1'z_1 \leqq (n-1)/n; \qquad = 0 \quad \text{otherwise.} \tag{3.13}$$

Then, the expression for f^* is, after a change of variable, given by

$$f^*(B, \bar{X}) = \int_{R^p} c^* h(z)|S|^{-\frac{1}{2}}\left\{1 - \frac{n}{n-1}(z-\bar{X})'S^{-1}(z-\bar{X})\right\}^{\frac{1}{2}(n-p-3)} \times I_n((z-\bar{X})'S^{-1}(z-\bar{X}))\,dz \tag{3.14}$$

where $B = \text{sqrt}(S)$.

As a particular case, let D be a set in R^p, suppose $\mathscr{L}(X) = \mathscr{N}_p(\mu, \Sigma)$, and consider the problem of finding a MVUE for the parametric function $\varphi_D(\mu, \Sigma) \equiv P_{\mu,\Sigma}(X_1 \in D)$. Then f^* in (3.14) is the MVUE of $\varphi_D(\mu, \Sigma)$ if h in (3.14) is taken to be the indicator function of the set D. For $p = 1$, this estimate was derived independently by Kolmogorov (1950), Lieberman and Resnikoff (1955), Barton (1961), and Sathe and Varde (1969).

With the result of the foregoing paragraph, the MVUE of the multivariate normal density

$$(2\pi)^{-\frac{1}{2}p}|\Sigma|^{-\frac{1}{2}}\exp\{-\tfrac{1}{2}(x-\mu)'\Sigma'(x-\mu)\} \tag{3.15}$$

at x can be had by differentiating f^* above (i.e., a set derivative with respect to D). The resulting MVUE is

$$c^*|S|^{-\frac{1}{2}}\left\{1 - \frac{n}{n-1}(x-\bar{X})'S^{-1}(x-\bar{X})\right\}^{\frac{1}{2}(n-p-3)} I_n((x-\bar{X})'S^{-1}(x-\bar{X})). \tag{3.16}$$

Here, $n \geqq p+2$ is required. The estimate (3.16) has been given previously by Ghurye and Olkin (1969). There, Ghurye and Olkin derive MVUE's of many nonstandard parametric functions for the multivariate normal and the Wishart distributions by methods different from ours. Many of their results can be derived by application of Theorem 2.2 and the discussion of this section.

Example 2. As another illustration of Theorem 2.2, an application to U-statistics (Fraser (1957)) is considered. Let $\mathscr{P} = \{F\colon F$ is a continuous distribution function$\}$ and let $\Theta = \mathscr{P}$. Let $X = (X_1, \cdots, X_n)'$ be a sample from some $F \in \mathscr{P}$. We are to find the MVUE of an estimable parametric function $\varphi(F)$ with unbiased estimate $f(X)$.

Let H be the group of transformations under the composition operation defined by $H = \{h\colon R \to R$ such that h is one-to-one, onto, continuous and strictly increasing$\}$. The family $\mathscr{P}$ is invariant under the group $G = \{g_h\colon R^n \to R^n$ by $g_h(x) = (h(x_1), \cdots, h(x_n))'$ and $h \in H\}$. As is well known (Halmos (1946)), the MVUE of $\varphi(F)$ is

$$f^*(x) = \frac{1}{n!} \sum_{\pi \in \Pi} f(x_{\pi(1)}, \cdots, x_{\pi(n)})$$

where Π is the set of permutations on $\{1, 2, \cdots, n\}$. This result is easily derived without Theorem 2.2. However, this result provides an interesting example where the group $\bar{G} = \{\bar{g}_h\colon \mathscr{P} \to \mathscr{P}$ such that $\bar{g}_h(F) = F \circ h^{-1}$, $h \in H\}$ does not act transitively on the parameter space $\mathscr{P}$.

REFERENCES

[1] BARTON, D. E. (1961). Unbiased estimation of a set of probabilities. *Biometrika* **48** 227–229.

[2] BASU, D. (1955). On statistics independent of a complete sufficient statistic. *Sankhyā* **15** 377–380.

[3] BASU, A. P. (1964). Estimates of reliability for some distributions useful in life testing. *Technometrics* **6** 215–219.

[4] FRASER, D. A. S. (1957). *Nonparametric Methods in Statistics*. Wiley, New York.

[5] GHURYE, S. and OLKIN, I. (1969). Unbiased estimation of some multivariate probability densities. *Ann. Math. Statist.* **40** 1261–1271.

[6] HALMOS, P. R. (1946). The theory of unbiased estimation. *Ann. Math. Statist.* **17** 34–43.

[7] KOLMOGOROV, A. N. (1950). Unbiased estimators. *Izv. Akad. Nauk SSSR. Ser. Mat.* **14** 303–326.

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

[9] LIEBERMAN, G. J. and RESNIKOFF, G. J. (1955). Sampling plans for inspection by variables. *J. Amer. Statist. Assoc.* **50** 457–516.

[10] NACHBIN, L. (1965). *The Haar Integral*. Van Nostrand, Princeton.

[11] NEYMANN, J. and SCOTT, E. L. (1960). Correction for bias introduced by a transformation of variables. *Ann. Math. Statist.* **31** 643–655.

[12] OLKIN, I. and PRATT, J. (1958). Unbiased estimation of certain correlation coefficients. *Ann. Math. Statist.* **29** 201–211.

[13] SATHE, Y. S. and VARDE, S. D. (1969). On minimum variance unbiased estimation of reliability. *Ann. Math. Statist.* **40** 710–714.

[14] SCHWARTZ, R. E. (1966). Properties of invariant multivariate tests. Unpublished thesis, Cornell Univ.

[15] TAMHANKAR, M. V. (1967). A characterization of normality. *Ann. Math. Statist.* **38** 1924–1927.

[16] TATE, R. F. (1959). Unbiased estimation: Functions of location and scale parameters. *Ann. Math. Statist.* **30** 341–366.

[17] WASHIO, Y., MORIMOTO, H. and IKEDA, N. (1956). Unbiased estimation based on sufficient statistics. *Bull. Math. Statist.* **6** 69–93.

[18] WIJSMAN, R. A. (1959). Applications of a certain representation of the Wishart matrix. *Ann. Math. Statist.* **30** 597–601.

[19] WIJSMAN, R. A. (1965). Cross-sections of orbits and their applications to densities of maximal invariants. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* Univ. of California. **1** 389–400.