FIDUCIAL THEORY AND INVARIANT ESTIMATION[1]

By R. B. Hora[2] and R. J. Buehler

University of Minnesota

1. Introduction. A class Ω of distributions may be called invariant under a group $\mathcal{G}$ of transformations of the sample space $\{x\}$ if Ω is closed under $\mathcal{G}$; that is, whenever x has a distribution belonging to Ω, so does gx for any $g \varepsilon \mathcal{G}$. In a fairly large body of decision theoretic literature, invariance enters into criteria for solutions, or is used as a tool. Fiducial theory is also intimately connected with invariance theory. In the present paper some links will be established between the decision theoretic and fiducial viewpoints.

Fisher (1934) gave fiducial distributions which he considered to be appropriate for location and scale parameters. These solutions were studied in detail by Pitman (1939), who showed, among other things, how certain "best" estimators could be defined in terms of expectations with respect to fiducial distributions. In the terminology of decision theory, Pitman's estimators are "best invariant" or "minimax invariant" estimators. Fraser's (1961a), (1961b) group-theoretic approach to fiducial theory, using Haar measures, is useful in providing a precise mathematical framework which is consistent with Fisher's in the case of location and scale parameters, and apparently in most other cases as well.

The present paper shows how certain of Pitman's results can be generalized using Fraser's theory. The results in Section 5 on "best" invariant estimators defined by means of fiducial distributions could actually be formulated and proved without reference to fiducial theory. For example, Blackwell and Girshick (1954), p. 314, express the best estimator of a location parameter in terms of the conditional expectation of the first observation x_1 given the differences $x_2 - x_1, \cdots, x_n - x_1$. Thus fiducial theory may be regarded as a convenient, but not essential, tool for obtaining desirable estimators.

In Section 2, assumptions similar to Fraser's on the class of distributions are spelled out in detail. The equivalence of the fiducial distribution to a posterior distribution is pointed out. Theorem 2.1 establishes the identity $E_f^x H = E_a^\omega H$ where H is an invariant function of the observations and parameters, E_f^x denotes expectation with respect to the fiducial distribution given x, and E_a^ω denotes conditional expectation given any value of the ancillary statistic. The theorem is not restricted to location and scale parameter cases.

Section 3 gives four examples of location and scale parameter families whose generality increases from the case of one location parameter to that of two location and two scale parameters. The latter includes the Behrens-Fisher problem

Received 1 June 1964; revised 5 September 1965.

[1] Research supported by National Science Foundation Grants G19126 and GP3816.

[2] Now at Purdue University, Indianapolis, Indiana.

as a special case. It is to be emphasized that an assumption of the existence of sufficient statistics is not needed at any point, nor need the variates be either independent or identically distributed. The group structure is the essential ingredient which makes these other assumptions unnecessary.

In Section 4 a definition is given of an "invariantly estimable function" $\psi(\omega)$, $\omega \varepsilon \Omega$. The concept (but not the term) can be found in Lehmann (1959), p. 243. It is shown that when ψ is invariantly estimable, there exists a group of transformations of $\{\psi\}$ which is homomorphic with $\mathcal{G}$. This result is used to define invariant functions of x and ω in terms of invariant estimators of ψ. From the decision theoretic viewpoint, invariant estimability of ψ implies the existence of an "admissible group" with respect to the decision (i.e., estimation) problem in the sense of Blackwell and Girshick.

In Section 5 we consider the definition of best invariant estimators of invariantly estimable functions in terms of fiducial expectations. A number of examples are given (in addition to Pitman's) wherein "best" corresponds to "minimum mean square error," and the estimator is given explicitly in terms of fiducial expectations.

In Section 6 we consider the relation between confidence and fiducial limits, showing in particular that limits obtained from the "derived fiducial" distribution of an invariantly estimable function correspond to confidence limits in the sense of Neyman.

We will complete the present section with brief mention of other related work. The unpublished Wald Lectures of Tukey (1958) contained remarks relating fiducial distributions and Haar measures, as well as a recognition of the fact that orbits in the sample space correspond to ancillary statistics. Other uses of Haar measures in related problems are found in Peisakoff (1950), Kudō (1955), Kiefer (1957), Wesler (1959), and Stone (1965).

Peisakoff (1950) identified parameters with elements of a transformation group, and obtained (under distributional assumptions not differing essentially from ours) theorems characterizing best invariant decision functions. Our results differ in making explicit use of the fiducial distribution and in the formulation of the nuisance parameter problem (Section 4 below).

General discussions of the invariant decision problem are given by Blackwell and Girshick (1954) and by Lehmann (1959). There is a considerable body of literature relevant to the question of whether best invariant procedures (such as those in Section 5 below) are minimax among the class of all procedures. We may mention Wald (1939), Hunt and Stein (see Lehmann (1959), Chapter 8), Girshick and Savage (1951), Peisakoff (1950), Kudō (1955), Kiefer (1957), and Wesler (1959). The last four give generalizations of the Hunt-Stein theorem. Kiefer (1957) gives a good review of earlier work in this general area.

Kiefer (1957) considered models which are more general than ours in several respects: randomized decisions and sequential problems are included, and the group which transforms Ω is not necessarily transitive on Ω. Kiefer shows, incidentally, that in the transitive case, it suffices for minimax considerations to restrict to nonrandomized decision functions.

The question of admissibility of Pitman estimators has been considered, for example, by Stein (1959b).

2. Generalization of Pitman's expectation identity. In this section we state our assumptions and generalize an identity due to Pitman (1939). Our distributional assumptions are essentially equivalent to those of Fraser (1961b), but the particular formulation is somewhat closer to that of Peisakoff (1950) or Stone (1965).

2.1 *Assumptions.*

ASSUMPTION 1. $(\mathfrak{X}, B_X)$, $(\mathfrak{T}, B_T)$, $(\mathfrak{A}, B_A)$ and (Ω, B_Ω) are measurable spaces such that there is a one-to-one correspondence between $\mathfrak{X}$ and $\mathfrak{T} \times \mathfrak{A}$,

$$x = (t, a), \tag{2.1}$$

and B_X corresponds to the minimal σ-field on $\mathfrak{T} \times \mathfrak{A}$ generated by B_T and B_A.

Here $\mathfrak{X}$ is the sample space, $a \varepsilon \mathfrak{A}$ is an ancillary statistic, and $t \varepsilon \mathfrak{T}$ is a conditionally sufficient statistic given the ancillary, and Ω is the parameter space.

ASSUMPTION 2. $\mathcal{G} = \{g\}$ is a group and $(\mathcal{G}, B_G)$ is a measurable space on which there exists a left invariant Haar measure μ satisfying

$$\mu(gG) = \mu(G) \quad \text{all} \quad g \varepsilon \mathcal{G}, \qquad G \varepsilon B_G. \tag{2.2}$$

We note that Assumption 2 will be satisfied whenever $\mathcal{G}$ is a locally compact topological group (Halmos (1950), Section 58).

ASSUMPTION 3. There exist one-to-one correspondences between the three spaces $\mathfrak{T}$, Ω and $\mathcal{G}$ such that all images of measurable sets are measurable.

Thus $\mathfrak{T}$ and Ω inherit a group operation from $\mathcal{G}$ and are isomorphic to $\mathcal{G}$.

ASSUMPTION 4. There is a family P^ω, $\omega \varepsilon \Omega$, of probability distributions on $\mathfrak{X}$ such that for corresponding $g \varepsilon \mathcal{G}$ and $\omega \varepsilon \Omega$,

$$P^\omega(X) = \int_X f(g^{-1}t \mid a)\lambda(da)\mu(dt) \quad \text{all} \quad X \varepsilon B_X, \tag{2.3}$$

where λ is a probability measure on $\mathfrak{A}$ and $f(\cdot \mid a)$ is a density with respect to μ on $\mathfrak{T}$ for each $a \varepsilon \mathfrak{A}$.

Here, following Fraser (1961b) and Stone (1965), we find it convenient to use the same symbol to denote corresponding points in different spaces. Thus in (2.3), $g^{-1}t$ denotes the element of $\mathfrak{T}$ corresponding to $g^{-1}t \varepsilon \mathcal{G}$ where $g \varepsilon \mathcal{G}$ is given, and $t \varepsilon \mathcal{G}$ corresponds to the integration variable $t \varepsilon \mathfrak{T}$. In all the following formulas the space to which a symbol belongs should be clear from the context.

Our final assumption rules out the possibility of different parameter values giving the same distribution. When this does occur, one would ordinarily arrange to identify equivalent values (see, for example, Kudō (1955) p. 33).

ASSUMPTION 5. If ω_1 and ω_2 are distinct, then P^{ω_1} and P^{ω_2} are not identical.

It is readily seen that the transformations of $\mathfrak{X}$ defined by $gx = (gt, a)$ form a group isomorphic to $\mathcal{G}$. At this point we might distinguish six spaces, all in one-to-one correspondence: $\mathcal{G}$, $\mathfrak{T}$, Ω, and the isomorphic groups which transform Ω, $\mathfrak{T}$, and $\mathfrak{X}$. The group $\mathcal{G}$ is exactly transitive on $\mathfrak{T}$ (i.e., $t_1, t_2 \varepsilon \mathfrak{T}$ imply a unique $g \varepsilon \mathcal{G}$ such that $t_1 = gt_2$) and on Ω, but not on $\mathfrak{X}$ unless $\mathfrak{A}$ contains but one point.

From (2.3) follows, incidentally,

$$P^{g\omega}(gX) = P^{\omega}(X) \quad \text{all} \quad g \,\varepsilon\, \mathcal{G}, \qquad \omega \,\varepsilon\, \Omega, \qquad X \,\varepsilon\, B_X . \tag{2.4}$$

The above assumptions are chosen for their convenience in our particular applications. From some points of view it might appear more natural to arrange the assumptions so as to relate more directly to the family P^{ω}. Thus one might begin with a group $\mathcal{G}$ of transformations of $\mathfrak{X}$, and a family P^{ω} closed under the group. From these and other assumptions, one might define $g\omega$ by means of the invariance condition (2.4), define the spaces $\mathfrak{A}$ and $\mathfrak{I}$ and derive results like (2.3). For approaches of this kind, the reader is referred to Kudō (1955), Fraser (1961b), and Hora (1964). When $\mathfrak{X}$ and $\mathcal{G}$ are taken as the starting point, then $\mathfrak{A}$ is defined as the space whose points are the *orbits* $\mathcal{G}x = \{gx \mid g \,\varepsilon\, \mathcal{G}\}$. The measure λ on $\mathfrak{A}$ is then induced from P^{ω}, and is shown not to depend on ω. The correspondence $x = (t, a)$ can be defined by assuming the existence of a "cross section" set $X \,\varepsilon\, B_X$ which contains exactly one point x_a of each orbit. The t coordinate of an arbitrary point x on the orbit having label a can be defined to correspond to that element $g \,\varepsilon\, \mathcal{G}$ for which $x = gx_a$. Topological and measure theoretic problems relating to the distributions on $\mathfrak{I}$ and $\mathfrak{A}$ have been considered by Kudō (1955) and Wijsman (1965).

2.2 *The fiducial distribution.* We define as usual a right Haar measure ν and a modular function Δ by

$$\nu(G) = \mu(G^{-1}), \qquad \mu(Gg) = \Delta(g)\mu(G) \quad \text{all} \quad g \,\varepsilon\, \mathcal{G}, \qquad G \,\varepsilon\, B_G . \tag{2.5}$$

From (2.2) and (2.5) follow

$$\mu(g(dt)) = \mu(dt), \qquad \mu((d\omega^{-1})t) = \Delta(t)\nu(d\omega). \tag{2.6}$$

Given the assumptions of Section 2.1, the fiducial distribution of ω given $x = (t, a)$ is defined to have the probability element

$$f(\omega^{-1}t \mid a)\Delta(t)\nu(d\omega). \tag{2.7}$$

The definition (2.7) is intended to agree with that of Fraser (1961a), (1961b), to whom the reader is referred for a discussion of its interpretation. A similar formula was mentioned in an incidental way by Peisakoff (1950), p. 38. For the special case of location and scale parameter families, (2.7) is consistent with the definitions of Fisher (1934) and Pitman (1939). Kudō (1955), p. 42, defines "the fiducial measure for z" (where z is the ancillary), which differs from (2.7) in that it depends only on the ancillary and not on the conditionally sufficient statistic t, and thus it gives the "shape" of the fiducial distribution without giving its "location."

As is fairly well known (Peisakoff (1950), p. 38, Wallace (1959), p. 872), (2.7) is equivalent to a posterior distribution when the prior measure is given by $\nu(\omega)$. This is evident because the likelihood can be factored into the marginal distribution of a times the conditional distribution of t given a. Since the former does not depend on ω, the posterior distribution is proportional to (2.7) and hence equal to it. Our examples are typical in that $\nu(\omega)$ is not a probability measure but an unbounded measure. Such prior measures are sometimes called "improper" distributions or "quasi" distributions (Wallace (1959), Stone (1965)). If the group $\mathcal{G}$ is compact so that the Haar measure is bounded, then the fiducial distribution is a posterior distribution for a true prior density. Such cases, while uncommon, can occur with rotation groups, as indicated in Section 3.2 below.

2.3 *The expectation identity.* We will denote expectation with respect to the fiducial distribution (2.7) by E_f^x (f denotes "fiducial"), and expectation with respect to the conditional distribution of x or t given a by E_a^ω. In the interest of economy these are preferred to the more complete notations: $E_f^{\omega \mid x}$ for E_f^x, and $E^{x \mid a,\omega}$ for E_a^ω.

THEOREM 2.1. *If the five assumptions of Section* 2.1 *are satisfied, and if* $H(x, \omega)$ satisfies

$$H(gx, g\omega) = H(x, \omega), \tag{2.8}$$

then

$$E_f^x H(x, \omega) = E_a^\omega H(x, \omega). \tag{2.9}$$

PROOF. Define $H(\cdot, \cdot, \cdot)$ by $H(t, a, \omega) = H(x, \omega)$. Then (2.8) gives

$$H(gt, a, g\omega) = H(t, a, \omega), \tag{2.10}$$

and by (2.6) and (2.10) (e denotes the identity element),

$$\begin{aligned} E_a^\omega H(x, \omega) &= \int H(t, a, \omega) f(\omega^{-1}t \mid a)\mu(dt) \\ &= \int H(\omega^{-1}t, a, e) f(\omega^{-1}t \mid a)\mu(\omega^{-1}(dt)) \\ &= \int H(s, a, e) f(s \mid a)\mu(ds) \\ &= \int H(\omega^{-1}t, a, e) f(\omega^{-1}t \mid a)\mu((d\omega^{-1})t) \\ &= \int H(t, a, \omega) f(\omega^{-1}t \mid a)\Delta(t)\nu(d\omega) \\ &= E_f^x H(x, \omega). \end{aligned}$$

2.4 *A counterexample.* We give an example to show that the expectation identity (2.9) can hold for a function H which does not satisfy the invariance condition (2.8). Take $p(x; \theta) = \frac{1}{4}$ for $\theta < x < \theta + 4$ and take

$$\begin{aligned} H(x, \theta) &= -(x + \theta), && \theta < x \leqq \theta + 1 \quad \text{or} \quad \theta + 3 < x \leqq \theta + 4 \\ &= +(x + \theta), && \theta + 1 < x \leqq \theta + 3 \\ &= 0, && \text{otherwise.} \end{aligned}$$

Here there is no ancillary and we may put E^ω for E_a^ω.

$$E^\omega H = \tfrac{1}{4}\{\textstyle\int_\theta^{\theta+1} + \int_{\theta+1}^{\theta+3} + \int_{\theta+3}^{\theta+4}\}H(x, \theta)\, dx$$

$$= \tfrac{1}{4}\{(-2\theta - \tfrac{1}{2}) + (4\theta + 4) + (-2\theta - 7/2)\} = 0,$$

$$E_f^x H = \tfrac{1}{4}\{\textstyle\int_{x-1}^{x} + \int_{x-3}^{x-1} + \int_{x-4}^{x-3}\}H(x, \theta)\, d\theta$$

$$= \tfrac{1}{4}\{(-2x + \tfrac{1}{2}) + (4x - 4) + (-2x + 7/2)\} = 0.$$

3. Examples of invariant families. In this section we give examples of families of distributions which satisfy the assumptions of Section 2. Four location and scale parameter families will be given in detail (Section 3.1) and some other cases will be mentioned briefly (Section 3.2).

3.1 *Details of four location and scale parameter families.* The general absolutely continuous bivariate distribution having a single location parameter θ has a density of the form $p(x_1 - \theta, x_2 - \theta)$. This class of distributions is easily seen to be equivalent to the class obtained by putting $x_1 = t$, $x_2 - x_1 = a$, where a has an arbitrary density $f(a)$ and the conditional density of t given a has the form $f(t - \theta \mid a)$. More generally the class $p(x_1 - \theta, \cdots, x_n - \theta)$ is obtained by giving an arbitrary density to the ancillary statistic represented by the $n - 1$ differences $(a_1, \cdots, a_{n-1}) = (x_2 - x_1, \cdots, x_n - x_1)$, and by taking the conditional density (given the ancillary) of $t - \theta = x_1 - \theta$ to be independent of θ but otherwise arbitrary. This n-dimensional case is our Example 3.1.

In general, location and scale parameters will be denoted by θ and σ respectively. Quantities x, y, θ, α range over $(-\infty, \infty)$ while σ, β range over $(0, \infty)$. Briefly we may describe the four cases thus:

EXAMPLE 3.1. θ;
EXAMPLE 3.2. (θ, σ);
EXAMPLE 3.3. $(\theta_1, \theta_2, \sigma)$;
EXAMPLE 3.4. $(\theta_1, \theta_2, \sigma_1, \sigma_2)$.

It may be noted that in none of the examples are the variates assumed to be independent and identically distributed (iid). Estimation of $\theta_1 - \theta_2$ in Example 3.4 is the Behrens-Fisher problem generalized to the non-normal, non-iid case.

The density for Example 3.4 has the form

$$(3.1)\quad \sigma_1^{-n}\sigma_2^{-m}p\{(x_1 - \theta_1)/\sigma_1, \cdots, (x_n - \theta_1)/\sigma_1, (y_1 - \theta_2)/\sigma_2, \cdots, (y_m - \theta_2)/\sigma_2\}.$$

Example 3.3 is obtained by putting $\sigma_1 = \sigma_2 = \sigma$; Example 3.2 is obtained by deleting the y variates and writing $(\theta_1, \sigma_1) = (\theta, \sigma)$; Example 3.1 is obtained from 3.2 by putting $\sigma = 1$. The space $(\mathfrak{X}, B_\mathfrak{X})$ is (R_n, B_n) in Examples 3.1 and 3.2 and (R_{n+m}, B_{n+m}) in Examples 3.3 and 3.4, where R_k is k-dimensional Euclidean space and B_k is the class of Borel sets. Table 3.1 gives the definitions of gx and $g\omega$. To save writing, gx is defined only on x_1 and y_1 with the understanding that the definition on other x's and y's is analogous. Table 3.2 gives the conditionally sufficient statistic t, where $t_1, \cdots, t_4$ are defined by

$$(3.2) \qquad t_1 = x_1, \quad t_2 = |x_1 - x_2|, \quad t_3 = y_1, \quad t_4 = |y_1 - y_2|,$$

and gives also the modular function Δ and the right Haar measure element $\nu(d\omega)$. The left Haar measure element $\mu(dg)$ is deducible from Δ and $\nu(d\omega)$ (in Example 3.4, $\mu(dg) = \beta_1^{-2}\beta_2^{-2}\, d\alpha_1\, d\alpha_2\, d\beta_1\, d\beta_2$). The definition of gt may be obtained by putting $(\theta, \sigma) = (\theta_1, \sigma_1)$ and by substituting (t_1, t_2, t_3, t_4) for $(\theta_1, \sigma_1, \theta_2, \sigma_2)$ respectively in the column $g\omega$ in Table 3.1.

Suitable representations of the ancillary statistic are: Example 3.1: $a = (a_1, \cdots, a_{n-1})$, $a_i = x_{i+1} - x_1$; Example 3.2: $a = (a_0, a_1, \cdots, a_{n-2})$, $a_i = (x_{i+2} - x_1)/(x_2 - x_1)$, $i = 1, \cdots, n-2$, $a_0 = \operatorname{sgn}(x_2 - x_1)(= -1 \text{ or } +1)$ according as $x_2 - x_1 < 0$ or $\geqq 0$); Example 3.3: $a = (a_0, a_1, \cdots, a_{n-2}, b_1, b_2, \cdots, b_{m-2}, c)$ with the a's as in Example 3.2, $b_j = (y_{j+2} - y_1)/(y_2 - y_1)$, $j = 1, \cdots, m-2$, $c = (y_2 - y_1)/(x_2 - x_1)$; Example 3.4: $a = (a_0, \cdots, a_{n-2}, b_0, \cdots, b_{m-2})$ where $b_0 = \operatorname{sgn}(y_2 - y_1)$ and the other quantities have the same definition as in Example 3.3.

TABLE 3.1

Example	ω	gx	$g\omega$
3.1	θ	$x_1 + \alpha$	$\theta + \alpha$
3.2	θ, σ	$\alpha + \beta x_1$	$\alpha + \beta\theta, \beta\sigma$
3.3	$\theta_1, \theta_2, \sigma$	$\alpha_1 + \beta x_1, \alpha_2 + \beta y_1$	$\alpha_1 + \beta\theta_1, \alpha_2 + \beta\theta_2, \beta\sigma$
3.4	$\theta_1, \sigma_1, \theta_2, \sigma_2$	$\alpha_1 + \beta_1 x_1, \alpha_2 + \beta_2 y_1$	$\alpha_1 + \beta_1\theta_1, \beta_1\sigma_1, \alpha_2 + \beta_2\theta_2, \beta_2\sigma_2$

TABLE 3.2

Example	t (see (3.2))	$\Delta(t)$	$\nu(d\omega)$
3.1	t_1	1	$d\theta$
3.2	t_1, t_2	t_2^{-1}	$\sigma^{-1}d\theta d\sigma$
3.3	t_1, t_2, t_3	t_2^{-1}	$\sigma^{-1}d\theta_1 d\theta_2 d\sigma$
3.4	t_1, t_2, t_3, t_4	$t_2^{-1}t_4^{-1}$	$\sigma_1^{-1}\sigma_2^{-1}d\theta_1 d\theta_2 d\sigma_1 d\sigma_2$

3.2 *Other examples of invariant families.* We briefly mention some other cases which are not considered in the later sections.

Clearly the location and scale parameter discussion could be extended to more than two σ's and more than two θ's in a straightforward manner. Stone (1965) gives this case as well as certain generalized and multivariate scale parameters.

For any bivariate distribution of variates (x, y) which is not symmetrical about the origin, one may consider the parametric family generated by rotation through an angle α about the origin. It is possible to obtain the fiducial distribution of α given n bivariate observations. Indeed the fiducial distribution equals the posterior distribution given a uniform prior over the interval $(0, 2\pi)$. The special case (x, y) independent normal with $Ex = R$, $Ey = 0$, $\operatorname{Var} x = \operatorname{Var} y = 1$,

$n = 1$, was considered by Fisher (1956), p. 135. More general cases have been considered by Hora (1964), particularly with regard to the problem of obtaining "best" estimators of α.

4. Invariant estimation. In this section we consider the estimation of a parameter point ω and of functions $\psi(\omega)$ which are required to be "invariantly estimable" according to a definition depending on the group structure assumed above. It is shown that for such functions ψ it is possible to define a group $\mathcal{G}'$ on $\{\psi\} = \Psi$ which is homomorphic to $\mathcal{G}$. It is further shown how invariant functions on $\mathfrak{X} \times \Omega$ can be defined in terms of invariant functions on $\mathfrak{X}$. In decision theoretic terms, where we identify the decision or action space with Ψ, the group $\mathcal{G}'$ corresponds to the group $\{g_A\}$ of Blackwell and Girshick (1954), p. 224, and an invariant loss function can be identified with a real-valued function of the above mentioned function on $\mathfrak{X} \times \Omega$. Thus invariant estimability of ψ allows us to define a $\mathcal{G}'$ to complete the triple $(\mathcal{G}, \mathcal{G}, \mathcal{G}')$, which transforms $(\mathfrak{X}, \Omega, \Psi)$, and which Blackwell and Girshick call an "admissible group" with respect to the decision problem. Finally it is shown how the space Ψ can be identified with a coset space, so that estimation of ψ is equivalent to the estimation of cosets.

4.1. *Estimation of* ω. Let $\hat{\omega}(x)$ be a mapping of $\mathfrak{X}$ onto Ω. We will say that $\hat{\omega}(x)$ is an invariant estimator of ω if

$$\hat{\omega}(gx) = g\hat{\omega}(x) \qquad \text{all } g \in \mathcal{G}. \tag{4.1}$$

This has been called the "principle of cogredience" by Lehmann (1950), Chapter 1, p. 17. Let us define

$$H(x, \omega) = \omega^{-1}\hat{\omega}(x). \tag{4.2}$$

If $\hat{\omega}(x)$ satisfies (4.1), then $H(gx, g\omega) = (g\omega)^{-1}\hat{\omega}(gx) = H(x, \omega)$, so that H satisfies (2.8).

4.2 *Invariantly estimable functions* $\psi(\omega)$. Frequently one does not wish to estimate the parameter point ω but only some function of it, say $\psi(\omega)$ with range Ψ. An equivalence relation "$\sim$" among elements of Ω is defined by

$$\omega_1 \sim \omega_2 \quad \text{means} \quad \psi(\omega_1) = \psi(\omega_2). \tag{4.3}$$

We will say that ψ is an *invariantly estimable function* (compare Lehmann (1959), p. 243) if

$$\omega_1 \sim \omega_2 \quad \text{implies} \quad g\omega_1 \sim g\omega_2 \quad \text{all} \quad g \in \mathcal{G}. \tag{4.4}$$

If ψ is a one-to-one function of Ω onto Ψ, then it satisfies (4.4) trivially. If not, a necessary and sufficient condition for ψ to be invariantly estimable is that $\psi(g\omega)$ have the form $\varphi(\psi(\omega))$. To illustrate, in Example 3.2, if $\psi(\omega) = \psi(\theta, \sigma) = \theta$, then $\psi(g\omega) = \alpha + \beta\theta$, which depends on ω only through $\psi(\omega) = \theta$, showing that ψ is invariantly estimable, as is also seen directly from (4.4). Similarly $\psi(\theta, \sigma) = \sigma$ is invariantly estimable, but θ/σ is not. Other examples which are easily verified are given in Table 4.1. Note that in Example 3.4 (generalized Behrens-Fisher case), the difference of means, $\theta_1 - \theta_2$, is not invariantly estimable. Fraser

(1961b), Section 12, has noted that in the Behrens-Fisher problem, "a fiducial interval for $\mu_1 - \mu_2$ will not be an invariant interval with respect to transformations for the x's and for the y's. In fact, under separate linear transformations for the x's and for the y's the parameter $\mu_1 - \mu_2$ is not transformed but is 'pulled apart'." The present section is intended to formalize and generalize Fraser's observation. Related remarks referring to interval estimation will be found in Section 6 below.

TABLE 4.1

Example	Invariantly Estimable	Not Invariantly Estimable
3.1	$\theta^{2n+1}, n = 0, 1, 2, \cdots$	$\theta^{2n}, n = 1, 2, \cdots$
3.2	$\theta, \sigma, c_1\theta + c_2\sigma$	$\theta/\sigma, \theta^3 + \sigma$
3.3	$\theta_1, \theta_2, \sigma,$ $c_1\theta_1 + c_2\theta_2 + c_3\sigma$	$\theta_1^3 + \theta_2^3, (\theta_1 - \theta_2)^2,$ $\theta_1\theta_2$
3.4	$\theta_1, \theta_2, \sigma_1, \sigma_2,$ $\sigma_1^r\sigma_2^s$	$\theta_1 \pm \theta_2, \sigma_1 \pm \sigma_2,$ $(\theta_1 + \sigma_1)/(\theta_2 + \sigma_2)$

4.3 *Invariant functions of* $\hat{\psi}$ *and* ω. For any point $\psi \varepsilon \Psi$ let ω denote any point of Ω such that $\psi(\omega) = \psi$. Then for any $g \varepsilon \mathcal{G}$ a transformation g' of Ψ onto Ψ is defined by

$$g'\psi(\omega) = \psi(g\omega), \tag{4.5}$$

and the definition is unique when (4.4) holds. In general, different elements g_1 and g_2 may define the same transformation g', e.g., in Example 3.2, if $\psi(\theta, \sigma) = \sigma$, then both $g_1 = (\alpha_1, \beta)$ and $g_2 = (\alpha_2, \beta)$ give $g'\sigma = \beta\sigma$. When g ang g' are in one-to-one correspondence, then there is an automatic isomorphism. Lemmas 4.1 and 4.2 below will be used to show that in any case a group operation can be defined on $\mathcal{G}' = \{g'\}$ such that the mapping of $\mathcal{G}$ onto $\mathcal{G}'$ is a homomorphism.

An equivalence relation "$\approx$" on $\mathcal{G}$ is defined by

$$g_1 \approx g_2 \quad \text{means} \quad g_1\omega \sim g_2\omega \quad \text{for all} \quad \omega \varepsilon \Omega. \tag{4.6}$$

LEMMA 4.1 *If* $g_1 \approx g_2$ *and* $g_3 \approx g_4$ *then* $g_1g_3 \approx g_2g_4$.

PROOF. Since $g_3 \approx g_4$, $\psi(g_3\omega) = \psi(g_4\omega)$; and using (4.4), $\psi(g_1g_3\omega) = \psi(g_1g_4\omega)$. Since $g_1 \approx g_2$, $\psi(g_1g_4\omega) = \psi(g_2g_4\omega)$. Thus $g_1g_3 \approx g_2g_4$.

LEMMA 4.2. *If* $g_1 \approx g_2$ *then* $g_1^{-1} \approx g_2^{-1}$.

PROOF. For any given g_1, g_2, ω, define $\omega' = g_2^{-1}\omega$ so that $\omega = g_2\omega'$. Since $g_1 \approx g_2$, $\psi(g_1\omega') = \psi(g_2\omega')$; and using (4.4), $\psi(g_1^{-1}g_1\omega') = \psi(g_1^{-1}g_2\omega')$. But also $\psi(g_1^{-1}g_1\omega') = \psi(g_2^{-1}g_2\omega')$ so that $\psi(g_1^{-1}g_2\omega') = \psi(g_2^{-1}g_2\omega')$, and substituting for ω' gives $\psi(g_1^{-1}\omega) = \psi(g_2^{-1}\omega)$.

Using Lemmas 4.1 and 4.2 we may now give unique definitions

$$g_1' \cdot g_2' = (g_1 \cdot g_2)' \quad \text{and} \quad (g')^{-1} = (g^{-1})', \tag{4.7}$$

and the mapping of $\mathcal{G}$ onto $\mathcal{G}'$ is a homomorphism.

The natural definition of an invariant estimator $\hat{\psi}(x)$ of ψ is

$$\hat{\psi}(gx) = g'\hat{\psi}(x). \tag{4.8}$$

We may note that a ψ-estimator defined in terms of an invariant ω-estimator is invariant; that is, it is easily shown that if $\hat{\omega}(x)$ satisfies (4.1) and $\hat{\psi}(x) = \psi(\hat{\omega}(x))$, then $\hat{\psi}$ satisfies (4.8).

Our next lemma gives the natural extension of the definition (4.2) which will be used in Section 5 below.

LEMMA 4.3. *Assume $\hat{\psi}(x)$ satisfies* (4.8), *let $\omega' \varepsilon \mathcal{G}'$ be the image of $\omega \varepsilon \Omega$, and define*

$$H(x, \omega) = (\omega')^{-1}\hat{\psi}(x). \tag{4.9}$$

Then H satisfies (2.8).

PROOF. $H(gx, g\omega) = ((g\omega)')^{-1}\hat{\psi}(gx) = (\omega')^{-1}(g')^{-1}g'\hat{\psi}(x) = H(x, \omega)$.

A real valued function of (4.9) will presently be identified with a loss function. Note that the right hand side of (4.9) depends on x only through $\hat{\psi}(x)$, $\psi \varepsilon \Psi$, and it depends on $\omega \varepsilon \Omega$ through the correspondence between ω and ω'. The space of decisions is identified with Ψ, so that the loss function is defined on $\Psi \times \Omega$ as desired.

4.4 *Remarks on subgroups and coset spaces.* Using the equivalence relation (4.3) we define

$$H = \{g \mid g \sim e\}, \qquad K = \{g \mid g\omega \sim \omega, \text{ all } \omega \varepsilon \Omega\}. \tag{4.10}$$

For all of the following incidental remarks, ψ is assumed to be invariantly estimable, so that equivalences can be left-multiplied (that is, $\omega_1 \sim \omega_2$ implies $g\omega_1 \sim g\omega_2$). Only one proof is given; the rest are left to the reader.

(i) $K \subset H \subset \mathcal{G}$. (ii) K is the kernel of the homomorphism between $\mathcal{G}$ and $\mathcal{G}'$, and it follows from group theory that K is a normal subgroup of $\mathcal{G}$ and that the quotient group $\mathcal{G}/K$, whose elements are the cosets gK, is isomorphic with $\mathcal{G}'$. (iii) H is a subgroup, but not necessarily a normal subgroup of $\mathcal{G}$. (iv) The cosets gH are in one-to-one correspondence with the values of ψ.

We indicate the proof of (iv), showing the equivalence of (a) $g_1 \sim g_2$, and (b) $g_1H = g_2H$. If (a) holds, then $g_2^{-1}g_1 \sim e$, so that $g_2^{-1}g_1 = h$, $(h \varepsilon H)$, whence $g_1 = g_2h$. Conversely if (b) holds, $g_1 = g_2h$, $(h \varepsilon H)$, whence $g_2^{-1}g_1 \sim e$, and $g_2 \sim g_1$.

Kudō (1955), p. 55, considered briefly the "estimation of cosets;" the result (iv) shows that coset estimation is equivalent to estimation of an invariantly estimable function. Apparently different relationships between decision spaces and coset spaces are given by Peisakoff (1950), pp. 44, 73.

Table 4.2 gives some examples of the subgroups H and K.

TABLE 4.2

Example	$\psi(\omega)$	H	K
3.1	θ	$\{(0)\}$	$K = H$
3.2	θ	$\{(0, \beta)\}$	$\{(0, 1)\}$
3.2	σ	$\{(\alpha, 1)\}$	$K = H$
3.2	$\theta + \sigma$	$\{(\alpha, \beta): \alpha + \beta = 1\}$	$\{(0, 1)\}$
3.3	$c_1\theta_1 + c_2\theta_2$	$\{(\alpha_1, \alpha_2, \beta): c_1\alpha_1 + c_2\alpha_2 = 0\}$	$\{(\alpha_1, \alpha_2, 1): c_1\alpha_1 + c_2\alpha_2 = 0\}$
3.4	$\sigma_1^r\sigma_2^s$	$\{(\alpha_1, \beta_1, \alpha_2, \beta_2): \beta_1^r\beta_2^s = 1\}$	$K = H$

5. Examples of best invariant estimators. It is well known that by restricting attention to invariant procedures, the decision problem is simplified by the fact that such procedures have constant risk (Peisakoff (1950), p. 26, Blackwell and Girshick (1954), Theorem 8.6.3, Kiefer (1957), p. 57, Wesler (1959), p. 4). Thus a "best" invariant procedure (when one exists), is simply one having the minimum risk. We now wish to apply Theorem 2.1 to express best invariant estimators in terms of fiducial expectations. When x is observed, the decision will be to estimate $\psi(\omega)$ by $\hat{\psi}(x)$. The corresponding loss is assumed to depend on $(\omega')^{-1}\hat{\psi}$ where $\omega' \varepsilon \mathcal{G}'$ corresponds to $\omega \varepsilon \Omega$. (We may note incidentally that $(\omega')^{-1}\hat{\psi}$ is not necessarily a unique function of $\psi(\omega)$ and $\hat{\psi}$.)

THEOREM 5.1. *With the structure assumed in Section 2, suppose that $\psi(\omega)$ is invariantly estimable, and that the loss when $\psi(\omega)$ is estimated by $\hat{\psi}$ has the form $\Phi((\omega')^{-1}\hat{\psi})$ where $\Phi(\cdot)$ is a real-valued function having domain Ψ. Suppose that for each x there is a unique value $\hat{\psi} = \hat{\psi}_0(x)$ which minimizes*

$$E_f^x\Phi((\omega')^{-1}\hat{\psi}). \tag{5.1}$$

Then $\hat{\psi}_0(x)$ minimizes the expected loss amongst all estimators $\hat{\psi}(x)$ which satisfy the invariance condition (4.8).

PROOF. It can be shown that $\hat{\psi}_0(x)$ satisfies (4.8). Let Φ and Φ_0 correspond respectively to $\hat{\psi}$ and $\hat{\psi}_0$. By definition of $\hat{\psi}_0$, $E_f^x(\Phi - \Phi_0) \geqq 0$ for all x. By Lemma 4.3, both $(\omega')^{-1}\hat{\psi}$ and $(\omega')^{-1}\hat{\psi}_0$ satisfy the invariance condition (2.8), and hence so does $(\Phi - \Phi_0)$. By Theorem 2.1, $E_a^\omega(\Phi - \Phi_0) = E_f^x(\Phi - \Phi_0) \geqq 0$ for all a, and taking expectation with respect to the distribution of a gives $E^\omega\Phi \geqq E^\omega\Phi_0$.

COROLLARY 5.1. *When Ψ is a subset of the real line, and when $\Phi((\omega')^{-1}\hat{\psi})$ has the form*

$$\Phi((\omega')^{-1}\hat{\psi}) = \varphi(\omega)(\hat{\psi} - \psi)^2 \tag{5.2}$$

where $\varphi(\omega) > 0$, then

$$\hat{\psi}_0(x) = E_f^x(\psi(\omega)\varphi(\omega))/E_f^x(\varphi(\omega)) \tag{5.3}$$

is the minimum mean square error invariant estimator of ψ, that is, it minimizes $E^\omega(\hat{\psi} - \psi)^2$.

PROOF. Clearly $E_f^x\{\varphi(\omega)(\hat{\psi} - \psi)^2\}$ is minimized when $\hat{\psi} = \hat{\psi}_0$ given by (5.3). Since $\varphi(\omega)$ is a constant for the operator E_a^ω,

$$\begin{aligned} \varphi(\omega)E_a^\omega(\hat{\psi}_0 - \psi)^2 &= E_f^x\{\varphi(\omega)(\hat{\psi}_0 - \psi)^2\} \\ &\leqq E_f^x\{\varphi(\omega)(\hat{\psi} - \psi)^2\} = \varphi(\omega)E_a^\omega(\hat{\psi} - \psi)^2. \end{aligned}$$

Thus $E_a^\omega(\hat{\psi}_0 - \psi)^2 \leqq E_a^\omega(\hat{\psi} \leqq \psi)^2$, and therefore $E^\omega(\hat{\psi}_0 - \psi)^2 \leqq E^\omega(\hat{\psi} - \psi)^2$.

Table 5.1 gives six examples of invariantly estimable functions $\psi(\omega)$ and the corresponding expressions for $g'\psi(\omega)$ and $(\omega')^{-1}\hat{\psi}$ implied by the definitions of g given in Table 3.1. If the quantity $\lambda = (\omega')^{-1}\hat{\psi}$ equals zero when $\psi = \hat{\psi}$ then reasonable loss functions are $|\lambda|$, λ^2, λ^4, etc. If $\lambda = 1$ when $\psi = \hat{\psi}$ then one may use $|\lambda - 1|$, $(\lambda - 1)^2$, $(\lambda - 1)^4$, etc. Theorem 5.1 would apply in any of these cases.

TABLE 5.1

Example	$\psi(\omega)$	$g'\psi(\omega) = \psi(g\omega)$	$\lambda = (\omega')^{-1}\hat{\psi}$
3.1	θ	$\psi + \alpha$	$\hat{\psi} - \psi$
3.2	θ	$\beta\psi + \alpha$	$(\hat{\psi} - \psi)/\sigma$
3.2	σ	$\beta\psi$	$\hat{\psi}/\psi$
3.2	$\theta + \sigma$	$\beta\psi + \alpha$	$(\hat{\psi} - \psi)/\sigma$
3.3	$c_1\theta_1 + c_2\theta_2$	$\beta\psi + c_1\alpha_1 + c_2\alpha_2$	$(\hat{\psi} - \psi)/\sigma$
3.4	$\sigma_1^r\sigma_2^s$	$\beta_1^r\beta_2^s\psi$	$\hat{\psi}/\psi$

Table 5.2 shows how Φ can be chosen in each case so that Corollary 5.1 can be applied to give the minimum mean square error invariant estimator exhibited in the last column. The first three of the six examples were considered by Pitman (1939). Of course it is not to be inferred that Corollary 5.1 would apply in any example. For instance, in Example 3.1 if we take $\psi(\omega) = \theta^3$ instead of θ, then ψ is still invariantly estimable, $g'\psi = (\psi^{\frac{1}{3}} + \alpha)^3$, $\lambda = (\omega')^{-1}\hat{\psi} = (\hat{\psi}^{\frac{1}{3}} - \psi^{\frac{1}{3}})^3$. Here the loss function $\lambda^{\frac{2}{3}}$ leads to $\hat{\psi}(x) = (E_f^x\psi^{\frac{1}{3}})^3$. Of course this is simply a translation of the solution obtained in Table 5.2; the point is that it cannot be called a minimum mean square error invariant estimator of $\psi = \theta^3$.

TABLE 5.2

Example	$\psi(\omega)$	$\Phi(\lambda)$	Minimum mean square error invariant estimator
3.1	θ	λ^2	$E_f^x\psi$
3.2	θ	λ^2	$E_f^x(\sigma^{-2}\psi)/E_f^x(\sigma^{-2})$
3.2	σ	$(\lambda - 1)^2$	$E_f^x(\psi^{-1})/E_f^x(\psi^{-2})$
3.2	$\theta + \sigma$	λ^2	$E_f^x(\sigma^{-2}\psi)/E_f^x(\sigma^{-2})$
3.3	$c_1\theta_1 + c_2\theta_2$	λ^2	$E_f^x(\sigma^{-2}\psi)/E_f^x(\sigma^{-2})$
3.4	$\sigma_1^r\sigma_2^s$	$(\lambda - 1)^2$	$E_f^x(\psi^{-1})/E_f^x(\psi^{-2})$

6. Some relationships between fiducial limits and confidence limits. In this section we consider the question of whether fiducial distributions can be used to obtain confidence limits or confidence regions in the sense of Neyman. Relationships with Bayesian theory are also indicated.

We begin by indicating the construction of invariant confidence sets. (Such sets are discussed by Lehmann (1959), Section 6.10, who considers optimal properties.) When $t = e$ (the identity), $\Delta(t) = \Delta(e) = 1$, and the fiducial probability element (2.7) becomes $f(\omega^{-1} \mid a)\nu(d\omega)$. Suppose that for each $a \in \mathfrak{A}$, a set $R_a \subset \Omega$ is determined which satisfies

$$\int_{R_a} f(\omega^{-1} \mid a)\nu(d\omega) = \gamma, \tag{6.1}$$

where $0 < \gamma < 1$. Let $I_e(\omega, a)$ be the indicator function which equals 1 for $\omega \in R_a$ and zero elsewhere. Defining further $I(x, \omega) = I_e(t^{-1}\omega, a)$, we have $I(gx, g\omega) = I(x, \omega)$. From (6.1) follows $E_f^xI(x, \omega) = \gamma$, all x, so that the region $S_x = \{\omega \mid I(x, \omega) = 1\}$ has fiducial probability γ for all x. By Theorem 2.1,

$$E^\omega I(x, \omega) = E^\omega E_a^\omega I(x, \omega) = E^\omega E_f^x I(x, \omega) = E^\omega\gamma = \gamma \tag{6.2}$$

for all ω, which shows that the system of regions $\{S_x\}$ are confidence regions in the sense of Neyman with confidence coefficient γ. Thus regions constructed in the above natural way from fiducial distributions (in the sense of Fraser) invariably have the frequency interpretation associated with confidence regions. A similar result for the location parameter case was pointed out by Pitman (1939), p. 396.

We now consider a different construction, also seemingly natural, wherein the confidence region property may be lost. Suppose that $\psi(\omega)$ is a real-valued measurable function, not necessarily invariantly estimable. The fiducial distribution of ω then defines a "derived" or "induced" distribution of ψ having percentile points $\bar{\psi}(x, \gamma)$ satisfying

$$(6.3) \qquad P_f\{\psi(\omega) \leqq \bar{\psi}(x, \gamma) \mid x\} = \gamma$$

where P_f denotes fiducial probability. In most cases of interest, $\bar{\psi}(x, \gamma)$ satisfying (6.3) will exist uniquely, and we will suppose for simplicity that this is the case. The "derived fiducial limits" $\bar{\psi}(x, \gamma)$ will be said to have the confidence interval property if

$$(6.4) \qquad P\{\psi(\omega) \leqq \bar{\psi}(x, \gamma) \mid \omega\} = \gamma \qquad \text{for all} \qquad \omega.$$

THEOREM 6.1. *If* (i) $\psi(\omega)$ *is invariantly estimable,* (ii) (6.3) *has a unique solution for* $\bar{\psi}(x, \gamma)$, *and* (iii) $g'\psi$ *increases as* ψ *increases for each* $g' \varepsilon \mathcal{G}'$, *then* $\bar{\psi}(x, \gamma)$ *has the confidence interval property.*

PROOF. The defining equation for $\bar{\psi}(x, \gamma)$ is

$$(6.5) \qquad \int_{\psi(\omega) \leqq \bar{\psi}(x,\gamma)} f(\omega^{-1}t \mid a)\Delta(t)\nu(d\omega) = \gamma.$$

Substituting $gx = (gt, a)$ for $x = (t, a)$ gives

$$(6.6) \qquad \int_{\psi(\omega) \leqq \bar{\psi}(gx,\gamma)} f(\omega^{-1}gt \mid a)\Delta(gt)\nu(d\omega) = \gamma.$$

Putting $\omega_1 = g^{-1}\omega$ and using the Haar measure identities $\Delta(gt) = \Delta(g)\Delta(t)$ and $\Delta(g)\nu(g(d\omega_1)) = \nu(d\omega_1)$ gives

$$(6.7) \qquad \int_{\psi(g\omega_1) \leqq \bar{\psi}(gx,\gamma)} f(\omega_1^{-1}t \mid a)\Delta(t)\nu(d\omega_1) = \gamma.$$

Since ψ is invariantly estimable, we may write $\psi(g\omega_1) = g'\psi(\omega_1)$, and comparison of (6.5) and (6.7) shows that $g'\psi(\omega) \leqq \bar{\psi}(gx, \gamma)$ is equivalent to $\psi(\omega) \leqq \bar{\psi}(x, \gamma)$. By assumption (iii) the first inequality is preserved under left multiplication by $(g')^{-1}$, so that by (ii) $(g')^{-1}\bar{\psi}(gx, \gamma) = \bar{\psi}(x, \gamma)$. It follows that the associated indicator function $I(x, \omega)$ (which equals 0 or 1 according as $\psi(\omega) >$ or $\leqq \bar{\psi}(x, \gamma)$) satisfies $I(gx, g\omega) = I(x, \omega)$, and the proof is completed as in (6.2).

The Behrens distribution of the difference δ of two normal means is known (from its relation to Bayesian analysis) to be a "derived fiducial" distribution in the above sense. Since δ is not invariantly estimable (Table 4.1), the known failure of Behrens' solution to have the confidence interval property might have been expected from the group theoretic viewpoint. Indeed, we conjecture that invariant estimability is a necessary condition for the conclusion of Theorem 6.1. For a different example exhibiting a more pronounced discrepancy, see Stein (1959a).

Finally, we may combine Theorem 6.1 with the relation noted in Section 2.2 between fiducial and posterior distributions.

COROLLARY 6.1. *When $\psi(\omega)$ is invariantly estimable, confidence limits obtained from the "derived fiducial" distribution correspond to Bayesian limits for an appropriate prior distribution.*

The special case corresponding to our Example 3.1 with $\psi(\omega) = \theta$ has previously been pointed out by Welch and Peers (1963), Section 5.

REFERENCES

BLACKWELL, D. and GIRSHICK, M. A. (1954). *Theory of Games and Statistical Decisions.* Wiley, New York.

FISHER, R. A. (1934). Two new properties of mathematical likelihood. *Proc. Roy. Soc. Ser. A* **14** 285–307.

FISHER, R. A. (1956). *Statistical Methods and Scientific Inference.* Oliver and Boyd, Edinburgh.

FRASER, D. A. S. (1961a). On fiducial inference. *Ann. Math. Statist.* **32** 661–676.

FRASER, D. A. S. (1961b). The fiducial method and invariance. *Biometrika* **48** 261–280.

GIRSHICK, M. A. and SAVAGE, L. J. (1951). Bayes and minimax estimates for quadratic loss function. *Proc. Second Berkeley Symp. Math. Statist. Prob.* 53–74. Univ. of California Press.

HALMOS, P. R. (1950). *Measure Theory.* Van Nostrand, Princeton.

HORA, R. B. (1964). Fiducial probability theory for distributions with a group structure. Ph.D. Thesis, and Technical Report Number 33, Department of Statistics, University of Minnesota.

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

KUDŌ, H. (1955). On minimax invariant estimates of the transformation parameter. *Natural Science Report of the Ochanomizu University* **6** 31–73.

LEHMANN, E. L. (1950). Notes on the theory of estimation (mimeographed). University of California, Berkeley.

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

PEISAKOFF, M. P. (1950). Transformation parameters. Ph.D. Thesis, Princeton University.

PITMAN, E. J. G. (1939). The estimation of the location and scale parameters of a continuous population of any given form. *Biometrika* **30** 391–421.

STEIN, C. (1959a). An example of a wide discrepancy between fiducial and confidence intervals. *Ann. Math. Statist.* **30** 877–880.

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

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

TUKEY, J. W. (1958). Fiducial inference. Inst. Math. Statist. Wald Lectures. Unpublished.

WALD, A. (1939). Contributions to the theory of statistical estimation and testing hypotheses. *Ann. Math. Statist.* **10** 299–326.

WALLACE, D. L. (1959). Conditional confidence level properties. *Ann. Math. Statist.* **30** 864–876.

WELCH, B. L. and PEERS, H. W. (1963). On formulae for confidence points based on integrals of weighted likelihoods. *J. Roy. Statist. Soc. Ser. B.* **25** 318–329.

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

WIJSMAN, R. A. (1965). Cross sections of orbits and their applications to densities of maximal invariants. Presented at Fifth Berkely Symposium on Mathematical Statistics and Probability.