Natural Science Report, Ochanomizu University, Vol. 6, No. 1, 1955
お茶の水女子大学 自然科学報告 第6巻 第1号

On Minimax Invariant Estimates of the Transformation Parameter

Hirokichi Kudō (工藤弘吉)

Department of Mathematics, Faculty of Science,
Ochanomizu University, Tokyo

Among decision problems in the statistical theory, there are many cases where the distribution space Ω remains invariant under any transformation belonging to a group G of transformations defined on the sample space. For example, if $(x_1, x_2, \ldots, x_n)$ is the sample of a fixed size from a normal distribution with unknown mean and unknown variance, then the normality of the distribution of the sample remains invariant under any transformation $(x_1, x_2, \ldots, x_n) \rightarrow (ax_1+b, ax_2+b, \ldots, ax_n+b)$. Of course, the parameters involved in the distribution function vary within Ω according as the transformation does. If an initial distribution is given in each type, the parameters can be considered as induced from the initial distribution under transformations on the sample space. Taking account of this fact, such a set of parameters should be called a transformation parameter. The pair of the location and scale parameters is an example of such parameters.

The group theoretic properties of transformations on the sample space were often used as a key of solving some special problems (*e.g.* Pitman [18, 19], Wolfowitz [23] and the author [13, 14]). In the general problem of estimation of the transformation parameter, are these properties powerful tools?

On the other hand, if the observations are our only source of information of the value of the transformation parameter, a satisfactory solution of the problem should remain invariant under any transformation of the group (the principle of invariance). Hence it seems desirable that the minimax solution of such a problem (if exists) is also invariant. Does it hold true?

The third question is: Can the invariant measure on Ω or a sequence of its truncated measures be used as a least favourable a priori distribution, relative to which the minimax invariant decision function is a Bayes solution? If the answer of this question is affirmative, the classical Bayes' method of solving the problem by using the uniform distribution as an a priori distribution becomes to have a reasonable foundation from the viewpoint of the subjective probability (see de Finetti [7]). In this connection Karlin [12] has proved a theorem on the theory of games that the invariant Haar measure on a compact group as a space of pure strategies is a minimax mixed strategy.

The above three questions are answered, in this paper, in the affirmative under certain conditions.

A curious one of these conditions is that the group $\overline{G}$ of transformations on Ω induced by G is an A-group (for definition, see page 45). But this condition seems to be unavoidable, since the special groups under which Hunt and Stein have proved the existence of the most stringent test are all A-groups (see Lehmann [16] and our Theorem 2.1)

The existence of the invariant measure on Ω is guaranteed by Mibu's theorem, as an extension of Haar measure on a locally compact group. Mibu [17] has developed his theory of the invariant measure on a locally compact and σ-compact uniform space admitting a group of transformations, and introduced a topology of the group such that the given space becomes a coset space. His theory gives us an important tool of our discussions. Hereupon we need to define the topology of Ω. Fortunately the absolute variation $V(P, P')$ of the difference of two distributions P and P' of Ω can be properly used as a metric of Ω, because $V(P, P')$ becomes larger according as the discrimination of P against P' by testing procedure is more difficult (see Kudō [13, 15]).

In the problem of estimation of the transformation parameter, it is not a considerable restriction to suppose that the decision space admits a group $\widetilde{G}$ isomorphic or homomorphic to $\overline{G}$, and that the loss function $W(P, \boldsymbol{a})$ remains invariant under any simultaneous and corresponding transformation $\sigma \in G$ of the distribution P and of the decision $\boldsymbol{a}$, *i.e.* $W(\bar{\sigma}P, \tilde{\sigma}\boldsymbol{a}) = W(P, \boldsymbol{a})$ for any $\sigma \in G$. For instance, when the decision space coincides with Ω, V is one of such functions, and hence every function on $\Omega \times \Omega$ only through V can be also used as a loss function. In this paper we restrict ourselves to consider the problem of estimation of the transformation parameter under a loss function mentioned above. However, for the generality of discussions, we speak of decision functions instead of estimates. By doing so, our results stand more close to Lehmann's theorem [16, Th. 9.1].

The concept of the invariant estimate was introduced by Fisher [8] at an early stage in the development of statistical theory, and has proved extremely fruitful. In 1938 Pitman [18] proposed a general method of constructing the invariant estimate of the location and scale parameters (simulaneously and separately) by using the notion of the "fiducial function". (See example *c*). The rearrangement of Pitman's estimate from the standpoint of the theory of decision functions was given by Girshick and Savage [9]. They have proved that Pitman's estimate is minimax with respect to the quadratic loss function $(\theta - a)^2$, where θ is the parameter and a is the estimated value. In Blackwell-Girshick's 1954 Book [4], the existence of the minimax invariant estimate of the location parameter is proved in the case of discrete variables in Euclidean space under a certain general condition of the loss function $W(\theta-a)$ which is satisfied either by every bounded $W(u)$ or by every continuous $W(u)$ having the infinite limits at $u=\pm\infty$. (See Example *b*). Discussions of the abstract transformation parameter problem have been already developed in 1950 by Hunt, Stein and Lehmann (see Lehmann [16]), but they concerned only the problem of tesing hypotheses.(0)

In the first section of this paper, we refer to Mibu's result without any proof and give the consideration about invariant statistics, which is inevitable in the following studies. Section 2 is the trunk of the paper. We shall give there a straightforward extension (Theorem 2.2) of Blackwell-Girshick's result [4, Theorem 11.3.1], which answers the above three questions. The problem of estimation of cosets, like that of estimation of the mean from the observation of normal variables with unknown variance and a problem of estimation by interval are discussed in the last of the section (Theorems 2.6, 2.7). It happens often that the sample space admitting a group G has an *a priori* topological structure. Section 3 is devoted to the proof of the homomorphism between two topological groups G and $\overline{G}$, when P is absolutely continuous with respect to the invariant measure on the sample space. The results obtained in Section 3 prove their worth in Section 4, in which some illustrative examples are gathered. The last example *f*) of Section 4 is given as an open problem.

1. Group of transformations of the sample space.

1.1. Topology of the parameter group. Mibu's Theorem.

Let $(X, \mathfrak{B}, P)$ be a probability space,(1) and G be a group of 1:1 measurable transformations(2) $e, \sigma, \tau, \rho, \ldots$ of X onto itself (e is the invariant transformation). Consider the probability measure $P(\sigma^{-1}B)$ on the measurable space $(X, \mathfrak{B})$, and denote it by $\bar{\sigma}P(B)$. For such measures $\bar{\sigma}P$, it holds that $\bar{\sigma}(\bar{\rho}P)=(\overline{\sigma\rho})P$, and we may write as $\overline{\sigma\rho}P$ instead of $(\overline{\sigma\rho})P$. Since

$$G_0=\{\sigma : \overline{\sigma\rho}P=\bar{\rho}P \text{ for every } \rho\in G\}$$

is a normal subgroup of G, the factor group $\overline{G}=G/G_0$ may be considered as a group of transformations of the space $\Omega=\{\bar{\sigma}P : \sigma\in G\}$. From this

(0) [24] reports that Peisakoff obtained some results concerning the invariant minimax decision function. They are perhaps very closely related to our task. I regret that they are unpublished.

(1) We assume that every σ-field of subsets of the spaces appearing in this paper contains every single point set, and is generated by a countable number of subsets of the space.

(2) If $B\in\mathfrak{B}$ implies $\sigma^{-1}B\in\mathfrak{B}$, then we call the transformation σ to be measurable.

fact, we shall regard $\bar{\sigma}, \bar{\tau}, \ldots$ operating on elements of Ω as an element of the group $\bar{G}$.

For the subject of this paper, it is convenient to use our terminologies as follows:

Ω= the distribution space,
X= the sample space,
$\bar{G}$= the parameter group,
an element of Ω= a distribution,
an element of X= a sample point,
an element of $\bar{G}$= a transformation parameter.

As we have seen in our previous papers [13, 15], the absolute variation

$$V(\bar{\sigma}P, \bar{\tau}P) = \sup_{B\in\mathfrak{B}} [\bar{\sigma}P(B) - \bar{\tau}P(B)] - \inf_{B\in\mathfrak{B}} [\bar{\sigma}P(B) - \bar{\tau}P(B)]$$

of the difference of two probability measures $\bar{\sigma}P$ and $\bar{\tau}P$ can be used as a "measure" of information of the experiment(3) $(\bar{\sigma}P, \bar{\tau}P)$. Hence it seems natural for the statistical discussion to use $V(\bar{\sigma}P, \bar{\tau}P)$ as the metric in Ω(4). Thus Ω is regarded as a topological space with the metric V in the following studies.

Assumption A. *The metric space Ω is locally compact and σ-compact.*(5) *Hence Ω is separable and complete.*

Remark 1.1. Owing to A. Berger [1], the condition of the separability of Ω is equivalent to that Ω is dominated by a σ-finite measure $\mathfrak{l}$:

$$\sigma P(B) = \int_B p(x:\sigma)\,\mathfrak{l}(dx), \quad B \in \mathfrak{B}.$$

The general theory of the topology of a group of transformations and of the invariant measure on a uniform space was developed by Yoshimichi Mibu [17]. However we shall cite here his results in a more restricted form.

Theorem 1.1. (Mibu) *Suppose that Ω is a metric space*(6) *satisfying Assumption A, and that $\bar{G}$ is a group of isometric transformations of Ω onto itself, with respect to which Ω is homogeneous.*(7) *Taking as a neighbourhood of the neutral element $\bar{e}$ of $\bar{G}$ the subsets*

---

(3) See Blackwell [3] and Kudō [15].

(4) We can see easily that $V(\sigma P, \tau P)$ is a $\bar{G}$-invariant metric: (1) $V(\bar{\sigma}P, \bar{\tau}P)=0$ with the equality when and only when $\bar{\sigma}P=\bar{\tau}P$, (2) $V(\sigma P, \bar{\tau}P)=V(\tau P, \dot{\sigma}P)$, (3) $V(\sigma P, \bar{\tau}P) \leqq V(\sigma P, \bar{\rho}P)+V(\rho P, \bar{\tau}P)$ and (4) $V(\bar{\rho}\bar{\tau}P, \bar{\rho}\bar{\sigma}P) = V(\bar{\tau}P, \bar{\sigma}P)$ for all $\bar{\rho}\in\bar{G}$.

(5) A topological space E is said to be σ-compact, if E is the sum of a countable number of compact subsets.

(6) It does not assumed in Theorems 1.1–1.4 that Ω is a distribution space. These theorems hold when Ω is an arbitrary metric space with a $\bar{G}$-invariant metric V.

(7) A space E is said to be homogeneous with respect to a group G of transformations of E if, given any two points of the space, there is an element of the group which transforms one into the other.

$$U(K, \varepsilon)=\{\bar{\sigma} : V(\overline{\sigma\rho}P, \bar{\rho}P)<\varepsilon \text{ for every } \bar{\rho}P \in K\},$$

where K is a compact subset of Ω and ε is a real positive number, $\overline{G}$ becomes a locally totally bounded topological group satisfying the first axiom of countability[8] *and the mapping $(\bar{\rho}P, \bar{\sigma})\to\overline{\sigma^{-1}\rho}P$ of $\Omega\times\overline{G}$ onto Ω is continuous.*

A necessary and sufficient condition of $\overline{G}$ to be totally bounded is that Ω is compact.

Assumption B.[9] *If $\overline{\sigma_1\rho}P$, $\overline{\sigma_2\rho}P$, $\ldots$ is a Cauchy sequence of points $\in\Omega$ for every $\bar{\rho}\in\overline{G}$, then there exists a $\bar{\sigma}_0(\in\overline{G})$ such that $\overline{\sigma_n\rho}P$ tends to $\overline{\sigma_0\rho}P$ in the sense of the metric V as $n\to\infty$ for every $\bar{\rho}\in\overline{G}$.*

Theorem 1.2. (Mibu). *Suppose, in addition to the hypotheses of Theorem 1.1, that $\overline{G}$ fulfills Assumption B. Then $\overline{G}$ is complete, locally compact and σ-compact. The compactness of Ω is a necessary and sufficient condition of the compactnesss of $\overline{G}$. Since $\overline{G}$ is metrisable* [5, Ch. IX, §3], *the separability of $\overline{G}$ follows from the σ-compactness of $\overline{G}$.*

The subgroup $\overline{H}$ consisting of the transformations under which P remains invariant is closed in the topological group $\overline{G}$ under Assumption A. In fact, let $\bar{\sigma}\notin\overline{H}$. Since $\varepsilon=V(\bar{\sigma}P, P)>0$, the neighbourhood $U(\{\bar{\sigma}P\}, \varepsilon/2)\,\bar{\sigma}$ of $\bar{\sigma}$ have no common element with $\overline{H}$, that is, $\overline{H}$ is closed in $\overline{G}$.

Thus we can treat the space $\overline{G}/\overline{H}$ of the left cosets $\bar{\sigma}\overline{H}$ as a topological space, in which the neighbourhoods of $\bar{\sigma}\overline{H}$ have a form $U(K, \varepsilon)\bar{\sigma}\overline{H}$. It is evident that the mapping $\alpha : \bar{\sigma}\overline{H}\to\bar{\sigma}P$ of $\overline{G}/\overline{H}$ onto Ω is 1:1, and moreover we have the following

Theorem 1.3. *The mapping α is topological; in the other words, Ω is homeomorphic to $\overline{G}/\overline{H}$ under the mapping α.*

---

(8) Since $\bar{\tau}U(\{P\}, \varepsilon/3)P=\{\rho P: V(\rho P, \bar{\tau}P)<\varepsilon/3\}$, and since K is compact, there is a finite number of elements $\bar{\tau}_1, \bar{\tau}_2, \ldots, \bar{\tau}_n$ of $\overline{G}$ such that

$$K\subset\bigcup_{i=1}^{n}\bar{\tau}U(\{P\}, \varepsilon/3)P.$$

Hence we have

$$\bigcap_{i=1}^{n}U(\{\bar{\tau}_iP\}, \varepsilon/3)=\bigcap_{i=1}^{n}\bar{\tau}_iU(\{P\}, \varepsilon/3)\tau_i^{-1}$$
$$\subset\bigcap_{i=1}^{n}\bigcap_{\sigma\in U(\{P\}, \varepsilon/3)}\bar{\tau}_i\bar{\sigma}U(\{P\}, \varepsilon)\bar{\sigma}^{-1}\tau_i^{-1}$$
$$=\{\bar{\rho} : V(\overline{\sigma\rho\tau_i}P, \overline{\tau_i\sigma}P)<\varepsilon \text{ for } \bar{\sigma}\in U(\{P\}, \varepsilon/3),\ i=1, 2, \cdots n\}$$
$$\subset\{\bar{\rho} : V(\overline{\rho\tau}P, \bar{\tau}P)<\varepsilon \text{ for every } \bar{\tau}P\in K\}$$
$$=U(K, \varepsilon).$$

This shows that $\{U(\{\tau P\}, r) : \tau\in\overline{G}, r=\text{a positive number}\}$ forms a complete system of neighbourhoods of the neutral element $\bar{e}$ in $\overline{G}$.

(9) If $\overline{G}$ is commutative or if G is the group of all isometric transformations on Ω, then this Assumption is superfluous.

Proof. It is evident that α is continuous, since the mapping $\bar{\sigma}\to\bar{\sigma}P$ is continuous (the definition of the topology of $\overline{G}$), and since the mapping $\bar{\sigma}\to\bar{\sigma}\overline{H}$ is open.

That α is open follows from the fact that

$$\alpha[U(\bar{\sigma}P, \varepsilon)\bar{\sigma}\overline{H}]=\{\bar{\tau}\bar{\sigma}P: V(\bar{\tau}\bar{\sigma}P, \bar{\sigma}P)>\varepsilon\}.$$

Corollary. If every transformation of $\overline{G}$, except for the neutral element, transforms P to the other point of Ω, then Ω is homeomorphic to $\overline{G}$.

Theorem 1.4. (Mibu). *Suppose that Ω and $\overline{G}$ satisfy the hypotheses of Theorem 1.2. There exists an outer measure ξ^* on Ω, which satisfies the following conditions:*

i) $\xi^*(\bar{\sigma}A)=\xi^*(A)$, $\bar{\sigma}\in\overline{G}$, $A\subset\Omega$.

ii) *Every Borel set is measurable.*(10)

iii) *For any subset A of Ω, there exists a Borel set B such that*

$$\xi^*(A)=\xi(B).$$

(we shall write as ξ the outer measure ξ^* of measurable set).

iv) *If A is compact, then $\xi(A)<\infty$.*

v) $\xi^*(\Omega)<\infty$ *if and only if Ω is compact.*

vi) *ξ is a continuous measure; i. e. there corresponds a neighbourhood U of $\bar{e}$ of $\overline{G}$ for every positive ε and every measurable subset A ($\xi(A)<\infty$) of Ω, such that*

$$\xi[(\bar{\sigma}A\cup A)-(\bar{\sigma}A\cap A)]<\varepsilon,\ \bar{\sigma}\in U,$$

vii) *If μ^* is a left invariant Haar measure on $\overline{G}$ and $f(\bar{\rho}P)$ is a ξ^*-measurable function on Ω, then $f(\overline{\sigma^{-1}\rho}P)$ is a $\xi^*\mu^*$-measurable function on $\Omega\times G$.*

viii) *The outer measure satisfying the conditions* i), ii) *and* iii) *is unique.*

ix) *It is a necessary and sufficient condition for a measure π, defined on the class of the Borel subsets of Ω, being absolutely continuous with respect to ξ that there corresponds a neighbourhood U of $\bar{e}$ of $\overline{G}$ for every $\varepsilon>0$ and every compact subset K of Ω, such that*

$$[\textit{absolute variation of } (\bar{\sigma}\pi-\pi) \textit{ on } K]<\varepsilon,\ \bar{\sigma}\in U,$$

where $\bar{\sigma}\pi(A)=\pi(\overline{\sigma^{-1}}A)$ for every Borel set $A\subset\Omega$.

x) *For two measurable subsets A and B of positive ξ-measure, there is an element $\bar{\tau}\in\overline{G}$ such that $\xi(\bar{\tau}A\cap B)>0$.*

---

(10) Under Assumption A, the concept of Borel sets is equivalent to that of Baire sets. See P. R. Halmos [10; p. 220].

This measure ξ is regarded as an invariant measure on $\overline{G}/\overline{H}$.

Since the closed subgroup $\overline{H}=\{\bar{\sigma}:\bar{\sigma}P=P\}$ of the locally compact group $\overline{G}$ is also locally compact, there exists a left invariant Haar measure ν on $\overline{H}$. For any Borel measurable function $f(\bar{\sigma})$ on $\overline{G}$, the integral

$$(1.1)\qquad \int_{\overline{G}/\overline{H}}\xi(d\overline{\sigma H})\int_{\overline{H}}f(\overline{\sigma\tau})\nu(d\bar{\tau})$$

is invariant under any trausformation $\bar{\sigma}\to\bar{\rho}\bar{\sigma}$. In fact, writing

$$f^*(\overline{\sigma H})=\int_{\overline{H}}f(\bar{\sigma}\bar{\tau})\,\nu\,(d\bar{\tau}),$$

we have

$$(1.2)\qquad \int_{\overline{G}/\overline{H}}f^*(\overline{\rho\sigma H})\xi(d\overline{\sigma H})=\int_{\overline{G}/\overline{H}}f^*(\overline{\sigma H})\,\xi\,(d\overline{\sigma H}).$$

(1.2) shows the $\overline{G}$-invariance of the integral (1.1). Therefore, by the uniqueness of the left invariant Haar measure on $\overline{G}$, we have

$$\int_{\overline{G}}f(\bar{\sigma})\mu(d\bar{\sigma})=\int_{\overline{G}/\overline{H}}\xi(d\overline{\sigma H})\int_{\overline{H}}f(\bar{\sigma}\bar{\tau})\,\nu\,(d\bar{\tau}),$$

that is, we may write without any confusion that

$$\int_{\overline{G}}f(\bar{\sigma})\,\mu\,(d\bar{\sigma})=\int_{\Omega}\xi\,(d\bar{\sigma}P)\int_{\overline{H}}f(\bar{\sigma}\bar{\tau})\,\nu\,(d\bar{\tau}).$$

We shall denote by $\mathfrak{L}$ the σ-field of all Borel subsets of the parameter group $\overline{G}$ throughout the remainder of this paper.

1.2. Invariant statistics.

Every measurable mapping $t(x)$ of the sample space $(X, \mathfrak{B})$ into another measurable space $(T, \mathfrak{T})$ induces a decomposition of X into disjoint measurable sets $X_t=\{x:t(x)=t\}$, $t\in T$. (See footnote (1)). Writing

$$(\bar{\sigma}P)_T(E)=\bar{\sigma}P(t^{-1}(E)) \text{ for } E\in\mathfrak{T},$$

$\bar{\sigma}P$ is resolved as

$$\bar{\sigma}P(B\frown t^{-1}(E))=\int_E\bar{\sigma}P(B:t)(\bar{\sigma}P)_T(dt)$$

for $B\in\mathfrak{B}$ and $E\in\mathfrak{T}$, where $\bar{\sigma}P(B:t)$ is a non-negative measurable function on T depending on $B\in\mathfrak{B}$ and on $\bar{\sigma}P$. In general, $\bar{\sigma}P(B:t)$ is not uniquely determined. However if we can choose $\bar{\sigma}P(B:t)$ such that $\bar{\sigma}P(B;t)$ coincides with $\bar{\sigma}P(B\frown X_t:t)$ for I_T-almost all $t\in T$ and for all $\bar{\sigma}P\in\Omega$, and if $\bar{\sigma}P(B\frown X_t:t)$ is a probability measure on $(X_t, \mathfrak{B}_t)$, $\mathfrak{B}_t=\{B\frown X_t:B\in\mathfrak{B}\}$, then we call $t(x)$ a statistic and say that $\bar{\sigma}P(B:t)$ is a conditional probability of B given that $t(x)=t$. The set of all $(\bar{\sigma}P)_T$ is denoted by Ω^T.

If the ranges of two statistics $t(x)$ and $t'(x)$ are in a 1:1 correspondence β as measurable spaces such that the set $\{x:\beta(t(x))\neq t'(x)\}$ is of $\mathfrak{l}$-measure zero, then we say that $t(x)$ is equivalent to $t'(x)$.

Let g be a subgroup of G, and T be the class of the subsets $gx=\{x':x'=\sigma x,\ \sigma\in g\}$ of X. If the class $\mathfrak{T}$ of the measurable subsets of T can be defined such that the mapping $t:x\to gx$ is a statistic, then a statistic, equivalent to $t(x)$, is said to be *g-invariant*, or, *i.e.*, $t(x)$ is g-invariant if and only if

$$x'=\sigma x \text{ and } \sigma\in g \text{ imply } t(x')=t(x) \ \mathfrak{l}\text{-almost everywhere on } X.$$

Hence we have, for a g-invariant statistic $t(x)$,

$$(\bar{\tau}\bar{\sigma}P)_T=(\bar{\sigma}P)_T,\quad \tau\in g \text{ and } \sigma\in G,$$

and

$$\bar{\tau}\bar{\sigma}P(B:t)=\bar{\sigma}P(\tau^{-1}B:t)\qquad (\bar{\sigma}P)_T\text{-almost everywhere on } T$$

for all $\tau\in g$.

When $t(x)$ is a statistic invariant under a normal subgroup g of G, we can consider a group G^T of transformations σ_T on T defined as follows:

$$t(\sigma gx)=t(g\sigma x)=\sigma_T t(x)\ \mathfrak{l}\text{-almost everywhere on } X,$$

that is, if $\sigma\in G$ then

$$\sigma_T t(x)=t(\sigma x)\ \mathfrak{l}\text{-almost everywhere on } X.$$

And we have

(1.3) $$(\bar{\sigma}P)_T=\bar{\sigma}_T P_T,\quad \sigma\in G\,;$$

(1.4) $$\bar{\sigma}P(B\frown t^{-1}(E))=\int_E \bar{\sigma}P(B:t)\bar{\sigma}_T P_T(dt),\quad \sigma\in G,$$

when $\bar{\sigma}_T P_T(E)=P_T(\sigma_T^{-1}E)$ for every $E\in\mathfrak{T}$. Hence it holds evidently that

(1.5) $$\bar{\sigma}P(B:t)=P(\sigma^{-1}\ B:\sigma_T^{-1}t)\ \ \sigma_T P_T\text{-almost everywhere on } T.$$

Denote

$$G_0^T=\{\sigma:\overline{\sigma_T\rho_T}P_T=\bar{\rho}_T P_T \text{ for every } \rho\in G\}\ ,$$

and

$$\bar{G}_0^T=\{\bar{\sigma}:\overline{\sigma_T\rho_T}P_T=\bar{\rho}_T P_T \text{ for every } \rho\in G\},$$

Since the mapping $\sigma\to\bar{\sigma}$ of G onto $\bar{G}$ is an algebraic homomorphism with a kernel G_0, and since G_0^T is a normal subgroup of G, the factor group $\overline{G^T}=G/G_0^T$ is algebraically isomorphic to $\bar{G}/\bar{G}_0^T$. On the other hand, since $\overline{G^T}$ can be regarded as a group of transformations of Ω^T, the absolute variation metric V_T on Ω^T reduces $\overline{G^T}$ to a topological group by Mibu's method stated in Theorem 1.1.

Theorem 1.5. *If both of Ω and Ω^T satisfy Assumptions A and B, then the algebraic homomorphism $\gamma:\bar{\sigma}\to\bar{\sigma}_T$ of $\bar{G}$ onto $\bar{G}^T$ is open and continuous.*

Proof. Let F^T be a compact subset of Ω^T, and $\overline{(\tau_1)_T}P_T, \overline{(\tau_2)_T}P_T, \dots, \overline{(\tau_n)_T}P_T$ be an $\varepsilon/3$-net in F^T. Choosing one point $\bar{\tau}_i P$ from each complete inverse image of $\overline{(\tau_i)_T}P_T$ under the mapping $\bar{\tau}P\to\bar{\tau}_T P_T$ of Ω onto Ω^T, $i=1, 2, \dots, n$, and denoting $K=\{\bar{\tau}_1P, \bar{\tau}_2P, \dots, \bar{\tau}_nP\}$, every element $\bar{\sigma}$ belonging to $U(K, \varepsilon/3)$ satisfies

$$\begin{aligned} &V_T(\overline{\sigma_T\tau_T}P_T, \bar{\tau}_TP_T)\\ &\leqq V_T(\overline{\sigma_T\tau_T}P_T, \overline{\sigma_T(\tau_i)_T}P_T)+V_T(\overline{\sigma_T(\tau_i)_T}P_T, \overline{(\tau_i)_T}P_T)\\ &\qquad\qquad\qquad\qquad +V_T(\overline{(\tau_i)_T}P_T, \bar{\tau}_TP_T)\\ &\leqq 2V_T(\bar{\tau}_TP_T, \overline{(\tau_i)_T}P_T)+V(\bar{\sigma}\bar{\tau}_iP, \bar{\tau}_iP)\\ &<\frac{2}{3}\varepsilon+\frac{1}{3}\varepsilon=\varepsilon \end{aligned}$$

for every $\bar{\tau}_TP_T\in F^T$. Hence $\bar{\sigma}_T$ belongs to the neighbourhood $U(F^T, \varepsilon)$ of the neutral element $\bar{e}_T$ of $\bar{G}^T$. Thus the continuity of γ is shown. It follows from Assumptions A and B that γ is an open mapping. (See Pontrjagin [20, Theorem 13]).

By Theorem 1.5 the kernel $\bar{G}_0^T$ of the mapping γ is closed, and the natural topology of $\bar{G}/\bar{G}_0^T$ is equivalent to the topology of $\bar{G}^T=G/G_0^T$, induced by the metric V_T.

Theorem 1.6 *Under the same hypotheses as Theorem 1.5, $\bar{G}^T$ is isomorphic to $\bar{G}/\bar{G}_0^T$ under the mapping $\sigma G_0^T\to\bar{\sigma}\bar{G}_0^T$ as a topological group.*

In the following we shall give two examples of invariant statistics, each of which will play an important rôle in the later discussions.

a) *A G-invariant statistic $z(x)$.* Let $(Z, \mathfrak{D})$ be a measurable space of the range of $z(x)$. Since $G\sigma=G$ for every $\sigma\in G$, every point z of Z remains invariant under any σ_Z:

$$\sigma_Z z(x)=z(\sigma x)=z(x)$$

for every $\sigma\in G$ and $\mathfrak{l}$-almost everywhere on X.

H)ence, from (1.3)–(1.5), we have

$$(1.6)\qquad \bar{\sigma}_ZP_Z=P_Z,$$

$$(1.7)\qquad \bar{\sigma}P(B\frown z^{-1}(D))=\int_D\bar{\sigma}P(B:z)P_Z(dz),\ D\in\mathfrak{D}.$$

and

$$(1.8)\qquad \bar{\sigma}P(B:z)=P(\sigma^{-1}B:z)\quad P_Z\text{-almost everywhere on } Z.$$

From (1.6), Ω^Z consists only of one point P_Z.

b) *A G_0-invariant statistic $y(x)$.* Denote by $(Y, \mathfrak{C})$ the measurable space of the range of $y(x)$, and let $z(y)$ be a G-invariant statistic on $(Y, \mathfrak{C})$ onto a measurable space $(Z, \mathfrak{D})$. We have, similarly to (1.6)−(1.8),

$$\bar{\sigma}_Z P_Z = P_Z \text{ for every } \sigma \in G,$$

$$\bar{\sigma}_Y P_Y(C \frown z^{-1}(D)) = \int_D \bar{\sigma}_Y P_Y(C:z)\bar{\sigma}_Z P_Z(dz), \quad C \in \mathfrak{C}, \; D \in \mathfrak{D},$$

and

$$\bar{\sigma}_Y P_Y(C:z) = P_Y(\sigma_Y^{-1} C:z) \text{ for every } \sigma \in G,$$
$$P_Z\text{-almost everywhere on } Z.$$

Moreover, $z_1(x)=z(y(x))$ is also a G-invariant statistic on X, and we have

$$\bar{\sigma}P(B \frown y^{-1}(C \frown z^{-1}(D)))$$
$$= \int_D P_Z(dz) \int_C \bar{\sigma}P(B:y, z)\bar{\sigma}_Y P_Y(dy:z),$$

where $\sigma P(B:y, z)$ is a conditional probability of B for given $z_1(x)=z$ and $y(x)=y$.

Especially, if $\tau \in G_0$, we have

(1.9) $$\bar{\sigma}P(\tau^{-1}B:y, z) = \bar{\sigma}P(B:y, z) \quad \mathfrak{l}_Y\text{-almost everywhere on } Y.$$

Let $x \in X$ be fixed for a moment, and write $I_x = \{\rho : \rho x = x\}$ and $gx = \{\rho x : \rho \in g\}$. The correspondence $G_0\rho x \to \rho I_x G_0$ of $\{G_0 \rho x : \rho \in G\}$ onto the left coset space $G/I_x G_0$ is 1:1. Since G_0 is a normal subgroup of each of G and $I_x G_0$, and since $\bar{I}_x = \{\bar{\rho} : \rho x = x\}$ is the factor group of $I_x G_0$ modulo G_0, the correspondence $\rho I_x G_0 \to \bar{\rho}\bar{I}_x$ of $G/I_x G_0$ onto $\bar{G}/\bar{I}_x$ is also 1:1. From this fact we may consider that $Y_z = \{y : z(y) = z\}$ coincides with the topological space $\bar{G}/\bar{I}_x$ for some $x \in X_z = \{x : z_1(x) = z\}$.

Assumption C. *There is a $\mathfrak{B}$-measurable subset X_0 of X such that*

1) *P_Z-almost all X_z cross with X_0 at only one point x_z,*

2) *$J_z = \{\bar{\rho} : \rho x_z = x_z\}$ is closed and compact in $\bar{G}$,*

3) *the σ-field $\mathfrak{C}_z = \{C \frown Y_z : C \in \mathfrak{C}\}$ contains every Borel subset of Y_z which is considered as a σ-compact and locally compact space $\bar{G}/\bar{J}_z$,*

4) *the exceptional set $N_z \subset Y_z$ where (1.9) does not hold is independent of $\bar{\sigma} \in \bar{G}$ and $\tau \in G_0$,*

5) *there is one and only one G_0-invariant measure of total measure 1 on every $X_y = \{x : y(x) = y\}$.*

From 4) and 5) of Assumption C, $\bar{\sigma}P(B:y, z)$ is an invariant measure on X_y, and hence independent of $\bar{\sigma}$. We may write this invariant measure by $\mathfrak{n}(B:y, z)$. Therefore we have

(1.10) $$\bar{\sigma}P(B \frown y^{-1}(C \frown z^{-1}(D)))$$
$$= \int_D P_Z(dz) \int_C \mathfrak{n}(B:y, z)\bar{\sigma}_Y P_Y(dy:z), \quad D \in \mathfrak{D}, \; C \in \mathfrak{C},$$

and

(1.11) $$\int_X f(x)\bar{\sigma}P(dx)=\int_Z P_Z(dz)\int_{Y_z} f_{\mathfrak{n}}(y, z)\bar{\sigma}_Y P_Y(dy:z)$$

for $\mathfrak{B}$-measurable function $f(x)$, where

$$f_{\mathfrak{n}}(y, z)=\int_{X_y} f(x)\,\mathfrak{n}(dx:y, z).$$

Theorem 1.7. *Under Assumptions A, B and C, every G_0-invariant statistic $y(x)$ is sufficient for Ω, and the mapping $\bar{\sigma}P\rightarrow\bar{\sigma}_Y P_Y$ of Ω onto Ω^Y is isometric. And further G^Y is isomorphic algebraically to $\overline{G}$ under the mapping $\sigma_Y\rightarrow\bar{\sigma}$.*

Proof. From (1.10) and (1.11) follows the sufficiency of $y(x)$.

Let $\{C^+, C^-\}$ be a disjoint Hahn decomposition of Y with respect to an additive set function $(\bar{\sigma}_Y P_Y-\bar{\tau}_Y P_Y)$. Since

$$\bar{\sigma}_Y P_Y(C\frown z^{-1}(D))-\bar{\tau}_Y P_Y(C\frown z^{-1}(D))$$
$$=\int_D\{\bar{\sigma}_Y P_Y(C:z)-\bar{\tau}_Y P_Y(C:z)\}P_Z(dz),\ C\in\mathfrak{C},\ D\in\mathfrak{D},$$

it holds for every $C\in\mathfrak{C}$ that

$$\bar{\sigma}_Y P_Y(C^+:z)-\bar{\tau}_Y P_Y(C^+:z)\geqq\bar{\sigma}_Y P_Y(C:z)-\bar{\tau}_Y P_Y(C:z)$$

P_Z-almost everywhere on Z.

On the other hand, writing $B^+=y^{-1}(C^+)$,

$$\bar{\sigma}_Y P_Y(C^+:z)-\bar{\tau}_Y P_Y(C^+:z)=\bar{\sigma}P(B^+:z)-\bar{\tau}P(B^+:z)$$
$$\geqq\int_{C^+}\mathfrak{n}(B:y, z)\{\bar{\sigma}_Y P_Y(dy:z)-\bar{\tau}_Y P_Y(dy:z)\}\geqq\bar{\sigma}P(B:z)-\bar{\tau}P(B:z)$$

for every $B\in\mathfrak{B}$ and for P_Z-almost all $z\in Z$. Therefore

$$\bar{\sigma}P(B^+)-\bar{\tau}P(B^+)=\int_Z\{\bar{\sigma}P(B^+;z)-\bar{\tau}P(B^+:z)\}P_Z(dz)$$
$$\geqq\int_Z\{\bar{\sigma}P(B:x)-\bar{\tau}P(B:z)\}P_Z(dz)=\bar{\sigma}P(B)-\bar{\tau}P(B)$$

for every $B\in\mathfrak{B}$.

Similarly we have also, by writing $B^-=y^{-1}(C^-)$,

$$\bar{\sigma}P(B^-)-\bar{\tau}P(B^-)\leqq\bar{\sigma}P(B)-\bar{\tau}P(B)\ \text{for every}\ B\in\mathfrak{B}.$$

This shows that $\{B^+, B^-\}$ is a Hahn decomposition of X with respect to $(\bar{\sigma}P-\bar{\tau}P)$. Hence we have

$$V(\bar{\sigma}P, \bar{\tau}P)=V_Y(\bar{\sigma}_Y P_Y, \bar{\tau}_Y P_Y).$$

Since G is a group of transformations on X, we have

$$\bigcap_{x\in X} I_x=\{e\},\ I_x=\{\rho:\rho x=x\},$$

and hence

$$\bigcap_{x\in X} I_x G_0=G_0.$$

This shows that each σ outside of G_0 is a transformation of a coset space $G/I_x G_0$ for some x, and that G^Y is isomorphic to $\overline{G}=G/G_0$.

Assumption D. *There are a G-invariant statistic and a G_0-invariant statistic.*

By Theorem 1.7 it is sufficient for statistical problems to treat only a G_0-invariant statistic instead of the sample itself. Hence, in the remainder of this section and in Section 2, we assume without any loss of generality that G is isomorphic to $\overline{G}$ as a topological group, and that sample space X is the union of the topological space X_z each of which is homogeneous with respect to some factor group of G. Hence we shall use, without any confusion, the same notation for elements of G, $\overline{G}$ and G^Y.

By 2) of Assumption C, there is a unique Haar measure $Q(L:z)$ of total measure 1 on J_z. Let $L \in \mathfrak{L}$. It is proved similarly as Theorem F of [10, page 281] that $Q(\tau^{-1}L\frown J_z:z)$ is Borel measurable in X_z as a function of $x=\tau x_z$. Write

$$\sigma\overline{P}(L:z)=\int_{X_z} Q(\tau^{-1}L\frown J_z:z)\sigma P(d(\tau x_z):z).$$

The probability measure $\sigma\overline{P}(L:z)$ thus defined on $(G, \mathfrak{L})$ satisfies obviously that

$$\sigma\overline{P}(L;z)=\overline{P}(\sigma^{-1}L:z).$$

We call $\overline{P}(L:z)$ the *fiducial measure for z*. Denoting

$$\bar{f}_z(\tau)=f(\tau x_z)$$

for a measurable function $f(x)$, we have

$$\int_X f(x)\sigma P(dx)=\int_Z P_Z(dz)\int_G \bar{f}_z(\sigma\tau)\overline{P}(d\tau:z). \tag{1.12}$$

In fact,

$$\begin{aligned} f(\tau x_z)&=\int_{J_z} f(\tau\rho x_z)Q(d\rho:z)=\int_G f(\tau\rho x_z)Q((d\rho)\frown J_z:z)\\ &=\int_G f(\rho x_z)Q(\tau^{-1}(d\rho)\frown J_z:z)\\ &=\int_G \bar{f}_z(\rho)Q(\tau^{-1}(d\rho)\frown J_z:z), \end{aligned}$$

and

$$\begin{aligned} \int_X f(x)\sigma P(dx)&=\int_Z P_Z(dz)\int_{X_z} f(x)\sigma P(dx:z)\\ &=\int_Z P_Z(dz)\int_{X_z} f(\tau x_z)\sigma P(d(\tau x_z):z) \end{aligned}$$

$$=\int_Z P_Z(dz)\int_{X_z}\sigma P(d(\tau x_z):z)\int_G \bar{f}_z(\rho)Q(\tau^{-1}(d\rho)\frown J_z:z).$$

Hence by Robins' theorem [21, Th. 2]

$$=\int_Z P_Z(dz)\int_G \bar{f}_z(\rho)\bar{\sigma}\bar{P}(d\rho:z)$$

$$=\int_Z P_Z(dz)\int_G \bar{f}_z(\sigma\rho)\bar{P}(d\rho:z).$$

2. Minimax invariant decision functions.

2.1 Simple decision problems admitting a group of transformations.

In Wald's theory [22] of statistical decision functions, the decision problem of fixed sample size is determined by the sample space X, the distribution space Ω, the decision space $\mathbf{A}$, the space Φ of decision functions and the loss function $W(P', \mathbf{a}): P' \in \Omega$ and $\mathbf{a} \in \mathbf{A}$. In this section, we assume that

a) There is a group G of transformations on the sample space X, under each transformation σ of which a fixed probability measure P on X is mapped into $\bar{\sigma}P(B)=P(\sigma^{-1}B)$.

b) The distribution space Ω is the set of all images $\bar{\sigma}P$ of a fixed P under transformations $\sigma \in G$. Hence $\bar{G}$ is the parameter group as defined in 1.1. The metric space Ω with V as a metric satisfies Assumptions A and B. From Theorem 1.1 $\bar{G}$ becomes a topological group. Denote by Λ and $\mathfrak{L}$ the classes of all Borel subsets of Ω and $\bar{G}$ respectively.

c) $(\mathbf{A}, \mathfrak{A})$ is a measurable space, and $\tilde{G}$ is a group of 1:1 $\mathfrak{A}$-measurable transformations $\tilde{\sigma}$ on $\mathbf{A}$, and $\bar{G}$ is homomorphic to $\tilde{G}$ under the mapping $\bar{\sigma} \to \tilde{\sigma}$.

d) $W(\bar{\rho}\bar{\sigma}P, \tilde{\rho}\mathbf{a})= W(\bar{\sigma}P, \mathbf{a})$ holds for every $\rho, \sigma \in G$ and for every $\mathbf{a} \in \mathbf{A}$. $W(\bar{\sigma}P, \mathbf{a})$ is a non-negative and $\Lambda \times \mathfrak{A}$-measurable function on $\Omega \times \mathbf{A}$. Write

$$w(\mathbf{a})= W(P, \mathbf{a}),\ \mathbf{a} \in \mathbf{A}.$$

e) Φ is the class of the probability measures $\varphi(A:x)$ on $(\mathbf{A}, \mathfrak{A})$ depending on x, such that $\varphi(\tilde{\sigma}A:\sigma x)$ is an $\mathfrak{L} \times \mathfrak{B}$-measurable function on $G \times X$ for any fixed $A \in \mathfrak{A}$. $\varphi \in \Phi$ is called a decision function.

f) Assumptions C and D are fulfiled by X and G.

A decision problem $(X, \Omega, \mathbf{A}, W, \Phi)$ satisfying the conditions a)—f) will be called a *simple* decision problem admitting the group G.

A decision function $\varphi(A:x)$ ($\in \Phi$) satisfying the condition

$$\varphi(\tilde{\sigma}A:\sigma x)=\varphi(A:x)$$

is said to be *invariant*. The probability measure

$$D(A:\varphi, \bar{\sigma}P)=\int_X \varphi(A:x)\bar{\sigma}P(dx)=\int_X \varphi(A:\sigma x)P(dx)$$

on $(\mathbf{A}, \mathfrak{A})$ is called a distribution of decision under φ when $\bar{\sigma}P$ is true. If φ is invariant, then

$$D(A:\varphi, \bar{\sigma}\bar{\rho}P)=D(\bar{\sigma}^{-1}A:\varphi, \bar{\rho}P).$$

If g is a subgroup of G, and if $\varphi(A:x)$ is a function of x only through a g-invariant statistics for any fixed A, then $D(A:\varphi, \bar{\sigma}P)$ may be regarded as a function of the right coset space $\overline{G}/\overline{g}$ for any fixed $A \in \mathfrak{A}$ and φ.

The risk function $r(\bar{\sigma}P, \varphi)$ is one of the essential concepts in Wald's theory. It is defined as

$$r(\bar{\sigma}P, \varphi)=\int_X \bar{\sigma}P(dx)\int_{\mathbf{A}} W(\bar{\sigma}P, \mathbf{a})\varphi(d\mathbf{a}:x). \tag{2.1}$$

By applying the definition of $\bar{\sigma}P$ and the condition d) to the integral (2.1) and then transforming $\mathbf{a}\to\bar{\sigma}\mathbf{a}$ and $x\to\sigma x$, we have

$$r(\bar{\sigma}P, \varphi)=\int_X P(dx)\int_{\mathbf{A}} w(\mathbf{a})\varphi(\bar{\sigma}d\mathbf{a}:\sigma x).$$

From (1.12),

$$r(\bar{\sigma}P, \varphi)=\int_Z P_Z(dz)\int_G \overline{P}(d_T:z)\int_{\mathbf{A}} w(\mathbf{a})\varphi_{\mathfrak{n}}(\bar{\sigma}d\mathbf{a}:\sigma_Y\tau_Y y_z), \qquad y_z=y(x_z), \tag{2.2}$$

where $z(x)$ is a G-invariant statistic, and

$$\varphi_{\mathfrak{n}}(A:y)=\int_{X_y} \varphi(A:x)\,\mathfrak{n}(dx:y, z).$$

By applying Robins' theorem [21, Th. 2] and Fubini's theorem, it follows from the condition e) that $r(\bar{\sigma}P, \varphi)$ is a $\mathfrak{L}$-measurable function on $\overline{G}$ for any fixed $\varphi \in \Phi$. Hence $r(\bar{\sigma}P, \varphi)$ is also Λ-measurable on Ω since the mapping $\bar{\sigma}\to\bar{\sigma}P$ is open and continuous.

For the completeness, we shall give some definitions in the theory of decision functions.

i) A probability measure π on (Ω, Λ) is called an a priori distribution, and the integral of the risk function

$$r^*(\pi, \varphi)=\int_\Omega r(\bar{\sigma}P, \varphi)\pi(d\bar{\sigma}P)$$

is called an average risk with respect to the a priori distribution π.

ii) If a decision function $\varphi_\pi \in \Phi$ is such that

$$r^*(\pi, \varphi_\pi)\leqq r^*(\pi, \varphi) \text{ for any } \varphi \in \Phi,$$

then φ_π is called a Bayes solution relative to π. If, in general, $\pi_1, \pi_2, \ldots$ is a sequence of a priori distributions, and φ_0 is a decision function $\in \Phi$ such that

$$\lim_{n\to\infty}\{r^*(\pi_n, \varphi_0)-\inf_{\varphi\in\Phi} r^*(\pi_n, \varphi)\}=0.$$

then φ_0 is called a Bayes solution in the wide sense relative to the sequence $\pi_1, \pi_2, \ldots$

iii) A decision function $\varphi_0 \in \Phi$ is said to be minimax, if

$$\sup_{\bar{\sigma}P\in\Omega} r(\bar{\sigma}P, \varphi)\geqq\sup_{\sigma P\in\Omega} r(\bar{\sigma}P, \varphi_0)$$

holds for every $\varphi \in \Phi$.

iv) An a priori distribution π_0 is said to be least favourable, if

$$\inf_{\varphi\in\Phi} r^*(\pi_0, \varphi)\geqq\inf_{\varphi\in\Phi} r^*(\pi, \varphi)$$

holds for every a priori distribution π.

v) A decision function $\varphi \in \Phi$ is said to be non-randomized, if there exists a measurable mapping $\mathbf{a}(x)$ of $(X, \mathfrak{B})$ into $(\mathbf{A}, \mathfrak{A})$ such that

$$\varphi(A:x)=\begin{cases}1, & \text{if } A \ni \mathbf{a}(x),\\ 0, & \text{otherwise.}\end{cases}$$

vi) A decision function φ_0 is admissible, if there is no $\varphi \in \Phi$ such that

$$r(\bar{\sigma}P, \varphi_0)\geqq r(\bar{\sigma}P, \varphi) \text{ for all } \bar{\sigma}P \in \Omega,$$

and

$$r(\bar{\sigma}P, \varphi_0)>r(\bar{\sigma}P, \varphi) \text{ for some } \bar{\sigma}P \in \Omega.$$

2.2. A-groups.

In this paragraph we shall give for the purpose of the later discussions a definition of A-groups, and some examples of them.

Denote by μ^r the right invariant Haar measure of G. We shall say that a locally compact and σ-compact group G is an *A-group,* if corresponding to any compact set $J\subset G$ and to a positive ε there is a compact set $K\subset G$ such that

$$\mu^r(K)>0 \text{ and } 1-\frac{\mu^r(K)}{\mu^r(K\cdot J^{-1})}<\varepsilon. \tag{2.3}$$

Theorem 2.1. *Each of the following group is an A-group:*

1) *The additive group of the integrals.*

2) *The additive group of the real numbers.*

3) *A compact group.*

4) *The direct product group of a finite number of A-groups.*

5) *A locally compact, connected and commutative group.*

6) *The product $G_1\cdot G_2$ of a closed normal subgroup G_1, and a closed subgroup G_2 (not necessarily normal) such that $G_1\frown G_2=\{e\}$, and that both of G_1 and G_2 are A-groups.*

Proof. 1)-4) is evident.

5) follows from Theorem 41 of [20] and 2)-4).

For 6), it is sufficient to prove that (2.3) holds for the product set $J=J_1\cdot J_2$ of two compact sets $J_1\subset G_1$ and $J_2\subset G_2$. Since G_2 is an A-group, we can choose a compact set $F_2\subset G_2$ such that

$$(1-\varepsilon)\mu_2(F_2\cdot J_2^{-1})<\mu_2(F_2),$$

where μ_2 is the right invariant Haar measure on G_2. Write

$$I=\bigcup_{\sigma\in F_2\cdot J_2^{-1}}\sigma J_1\sigma^{-1}.$$

This set I, being the image of the compact subset $J_1\times(F_2\cdot J_2^{-1})$ of $G_1\times G_2$ by the continuous mapping $(\sigma,\tau)\to\sigma\tau\sigma^{-1}$, is also compact in the A-group G_1. Hence there is a compact subset F_1 of G_1, such that

$$(1-\varepsilon)\mu_1(F_1\cdot I^{-1})<\mu_1(F_1),$$

where μ_1 is the right invariant Haar measure on G_1. By writing

$$F=F_1\cdot F_2,$$

we have

$$\begin{aligned}\mu^r(F\cdot J^{-1})&=\mu^r(F_1\cdot F_2\cdot J_2^{-1}\cdot J_1^{-1})\\&=\int_{F_2\cdot J_2^{-1}}\mu_1(F_1\cdot\sigma\cdot J_1^{-1}\cdot\sigma^{-1})\mu_2(d\sigma)\\&\leqq\mu_1(F_1\cdot I^{-1})\mu_2(F_2\cdot J_2^{-1})\\&<(1-\varepsilon)^{-2}\mu_1(F_1)\mu_2(F_2)\\&=(1-\varepsilon)^{-2}\mu^r(F),\end{aligned}$$

since it holds for every $\sigma\in F_2\cdot J_2^{-1}$ that

$$F_1\cdot F_2\cdot J_2^{-1}\cdot J_1^{-1}\sigma^{-1}\frown G_1=F_1\cdot\sigma\cdot J_1^{-1}\cdot\sigma^{-1}.$$

In general, the question whether every σ-compact and locally compact group is an A-group remains open.

2.3 An extension of Blackwell-Girshick's theorem.

Similarly to the last part of the preceding section, we shall assume without any loss of generality that $G=\overline{G}=G^r$. (see page 42).

Fix a non-descending sequence $J_1 J_2,\ldots$ of compact subsets of G, tending to G, and write

$$\mathbf{b}_n(\mathbf{a},z)=\int_{J_n}w(\tilde{\rho}\mathbf{a})\overline{P}(d\rho:z),\ n=1,2,\ldots,$$

and

$$\mathbf{b}(\mathbf{a},z)=\int_G w(\rho\mathbf{a})\overline{P}(d\rho:z),$$

where $\overline{P}(L:z)$ is the fiducial measure for z.

Theorem 2.2. *Suppose that* $(X, \Omega, \mathbf{A}, W, \Phi)$ *is a simple decision problem admitting an A-group G. It holds for every decision function* $\varphi \in \Phi$ *that*

$$\sup_{\sigma P \in \Omega} r(\sigma P, \varphi) \geqq F^*,$$

if there is an integer n for any positive ε *and for any* $z \in Z-N$ $(P_Z(N)=0)$ *such that*

(2.4) $$\mathbf{b}_n(\mathbf{a}, z) > F(z) - \varepsilon \quad for\ every\ \mathbf{a} \in \mathbf{A},$$

where

$$F(z) = \inf_{a \in A} \mathbf{b}(\mathbf{a}, z) \text{ and } F^* = \int_Z F(z) P_Z(dz).$$

Proof. We assume without any loss of generality that

$$\sup_{\sigma P \in \Omega} r(\sigma P, \varphi) < \infty.$$

It follows from (2.2) and Fubini's theorem that

(2.5) $$\begin{aligned} R(K, \varphi) &= \int_K r(\sigma P, \varphi) \mu^r(d\sigma) \\ &= \int_K \mu^r(d\sigma) \int_Z P_Z(dz) \int_G \overline{P}(d\rho : z) \int_{\mathbf{A}} w(\mathbf{a}) \varphi(\tilde{\sigma} d\mathbf{a} : \sigma\rho x_z) \\ &= \int_Z P_Z(dz) \int_G \overline{P}(d\rho : z) \int_K \mu^r(d\sigma) \int_{\mathbf{A}} w(\mathbf{a}) \varphi(\tilde{\sigma} d\mathbf{a} : \sigma\rho x_z) \end{aligned}$$

for every compact set $K \subset G$. By applying the transformation $\boldsymbol{a} \to \tilde{\rho}\boldsymbol{a}$, we have

(2.6) $$\int_{\mathbf{A}} w(\mathbf{a}) \varphi(\tilde{\sigma} d\mathbf{a} : \sigma\rho x_z) = \int_{\mathbf{A}} w(\tilde{\rho}\mathbf{a}) \varphi(\tilde{\sigma}\tilde{\rho} d\mathbf{a} : \sigma\rho x_z).$$

Again by transforming $\sigma \to \sigma\rho^{-1}$,

(2.7) $$\int_K \varphi(\tilde{\sigma}\tilde{\rho} A' : \sigma\rho x_z) \mu^r(d\sigma) = \int_{K\rho} \varphi(\tilde{\sigma} A' : \sigma x_z) \mu^r(d\sigma)$$

for every $A' \in \mathfrak{A}$. From (2.5), (2.6) and (2.7), and by using Robins' theorem [21, Th. 2], we have

(2.8) $$R(K, \varphi) = \int_Z P_Z(dz) \int_G \overline{P}(d\rho : z) \int_{K\cdot\rho} \mu^r(d\sigma) \int_{\mathbf{A}} w(\tilde{\rho}\mathbf{a}) \varphi(\tilde{\sigma} d\mathbf{a} : \sigma x_z).$$

On the other hand, by Fubini's theorem and by the fact $\{(\rho, \sigma) : \rho \in G, \sigma \in K\rho\} = \{(\rho, \sigma) : \rho \in K^{-1}\cdot\sigma, \sigma \in G\}$,

$$\int_G \overline{P}(d\rho : z) \int_{K\cdot\rho} f(\sigma, \rho) \mu^r(d\sigma) = \int_G \mu^r(d\sigma) \int_{K^{-1}\cdot\sigma} f(\sigma, \rho) \overline{P}(d\rho : z)$$

holds P_Z-almost everywhere on Z for every $\mathfrak{L} \times \mathfrak{L}$-measurable function $f(\sigma, \rho)$. Since

$$\int_{\mathbf{A}} w(\tilde{\rho}\mathbf{a})\varphi(\tilde{\sigma}d\mathbf{a}:x_z)$$

is $\mathfrak{L}\times\mathfrak{L}$-measurable, (2.8) becomes

$$\int_{Z} P_Z(dz)\int_{G}\mu^r(d\sigma)\int_{K^{-1}\cdot\sigma}\overline{P}(d\rho:z)\int_{\mathbf{A}} w(\tilde{\rho}\mathbf{a})\varphi(\tilde{\sigma}d\mathbf{a}:\sigma x_z)\,.$$

Therefore from Fubini's theorem follows

$$R(K,\varphi)=\int_{Z} P_Z(dz)\int_{G}\mu^r(d\sigma)\int_{\mathbf{A}}\varphi(\tilde{\sigma}d\mathbf{a}:\sigma x_z)\int_{K^{-1}\cdot\sigma} w(\tilde{\rho}\mathbf{a})\overline{P}(d\rho:z)\,. \tag{2.9}$$

Now let ε be any arbitrary positive number, D_n a $\mathfrak{D}$-measurable set of all $z\in Z$ for which (2.4) holds, and K_n a compact set satisfying (2.3) for $J=J_n$. Writing

$$K_n^*=K_n\cdot J_n^{-1}\,, \tag{2.10}$$

K_n^* is compact and $K_n^{*-1}\cdot\sigma\supset J_n$ holds for every $\sigma\in K_n$. Hence (2.9) yields

$$\begin{aligned} R(K_n^*,\varphi)&\geqq\int_{D_n} P_Z(dz)\int_{K_n}\mu^r(d\sigma)\int_{\mathbf{A}}\mathbf{b}_n(\mathbf{a},z)\varphi(\tilde{\sigma}d\mathbf{a}:\sigma x_z)\\ &>\int_{D_n} P_Z(dz)\int_{K_n}\mu^r(d\sigma)\int_{\mathbf{A}}[F(z)-\varepsilon]\varphi(\sigma d\mathbf{a}:\sigma y_z)\\ &=\mu^r(K_n)\int_{D_n}[F(z)-\varepsilon]P_Z(dz)\\ &>(1-\varepsilon)\mu^r(K_n^*)\int_{D_n}[F(z)-\varepsilon]P_Z(dz). \end{aligned} \tag{2.11}$$

Evidently $\{D_n\}$ is a non-descending sequence and tends to $Z-N$, since $\{J_n\}$ is non-descending. Hence

$$\lim_{n\to\infty}\int_{D_n} F(z)P_Z(dz)=\int_Z F(z)P_Z(dz)=F^*\,.$$

That is to say, there is an integer n_0 such that

$$\int_{D_n} F(z)P_Z(dz)<F^*-\varepsilon \quad\text{for every}\quad n>n_0\,. \tag{2.12}$$

Hence we have, from (2.11) and (2.12),

$$\sup_{\sigma P\ni\Omega} r(\sigma P,\varphi)\geqq\frac{R(K_n^*,\varphi)}{\mu^r(K_n^*)}>(1-\varepsilon)(F^*-2\varepsilon), \tag{2.13}$$

where n is a larger integer than n_0.

Since ε is arbitrary, (2.13) shows that

$$\sup_{\sigma P\in\Omega} r(\sigma P,\varphi)\geqq F^*\,.$$

Thus the proof is complete.

This theorem is a straightforward extension of Blackwell-Girshick's result, and the proof developed above has been proceeded exactly along the way of theirs.

Corollary. Under the hypotheses of Theorem 2.2, the invariant decision function φ^0 is minimax, if and only if

(2.14) $$\varphi^0(A_z : x_z)=1 \quad P_Z\text{-almost everywhere on } Z,$$

provided that $A_z=\{\mathbf{a}:\mathbf{b}(\mathbf{a}, z)=F(z)\}$ is not empty.

Proof. For φ^0 satisfying (2.14) it follows from (2.2) and Fubini's theorem that

(2.15) $$\begin{aligned} r(\sigma P, \varphi^0) &= \int_Z P_Z(dz)\int_G \overline{P}(d\tau : z)\int_{\mathbf{A}} w(\mathbf{a})\varphi^0(\tilde{\sigma} d\mathbf{a} : \sigma_T x_z) \\ &= \int_Z P_Z(dz)\int_G \overline{P}(d\tau : z)\int_{\mathbf{A}} w(\tilde{\tau}\mathbf{a})\varphi^0(d\mathbf{a} : x_z) \\ &= \int_Z P_Z(dz)\int_{\mathbf{A}} \mathbf{b}(\mathbf{a}, z)\varphi^0(d\mathbf{a} : x_z) \\ &= \int_Z P_Z(dz)\int_{A_z} \mathbf{b}(\mathbf{a}, z)\varphi^0(d\mathbf{a} : x_z) \\ &= \int_Z P_Z(dz)\int_{A_z} F(z)\varphi^0(d\mathbf{a} : x_z) \\ &= \int_Z F(z)P_Z(dz)=F^*. \end{aligned}$$

Hence φ^0 is minimax by Theorem 2.2.

Denoting by K_n a compact subset of G, satisfying (2.3) for $J=J_n$, we have, similarly to (2.11),

$$\begin{aligned} \frac{R(K_n\cdot J_n^{-1}, \varphi)}{\mu^r(K_n\cdot J_n^{-1})} &\geqq \frac{\mu^r(K_n)}{\mu^r(K_n\cdot J_n^{-1})}\int_Z P_Z(dz)\int_{\mathbf{A}} \mathbf{b}_n(\mathbf{a}, z)\varphi(d\mathbf{a} : x_z) \\ &> (1-\varepsilon)\left[F^* + \int_Z P_Z(dz)\int_{\mathbf{A}}\left\{\mathbf{b}_n(\mathbf{a}, z)-F(z)\right\}\varphi(d\mathbf{a} : x_z)\right] \end{aligned}$$

for every invariant decision function φ. If φ is minimax, then

$$F^*=\sup_{\sigma P\in\Omega} r(\sigma P, \varphi) \geqq \frac{R(K_n\cdot J_n^{-1}, \varphi)}{\mu^r(K_n\cdot J_n^{-1})}.$$

Hence for a minimax and invariant φ we have

$$\frac{\varepsilon}{1-\varepsilon}F^* > \int_Z P_Z(dz)\int_{\mathbf{A}}\{\mathbf{b}_n(\mathbf{a}, z)-F(z)\}\,\varphi(d\mathbf{a} : x_z),$$

$$(n=1, 2, \ldots \text{ ad. inf.})$$

Since $\lim \mathbf{b}_n(\mathbf{a}, z)=\mathbf{b}(\mathbf{a}, z)\geqq F(z)$, it follows that

$$\int_Z P_Z(dz)\int_{\mathbf{A}}\{\mathbf{b}(\mathbf{a}, z)-F(z)\}\varphi(d\mathbf{a}:x_z)=0,$$

and

$$\int_{\mathbf{A}}\{\mathbf{b}(\boldsymbol{a}, z)-F(z)\}\varphi(d\boldsymbol{a}:x_z)=0 \quad P_Z\text{-almost everywhere on } Z.$$

Hence we have

$$\varphi(A_z, x_z)=1 \quad P_Z\text{-almost everywhere on } Z.$$

This Collolary shows that the minimax invariant decision function can be constructed if and only if $\mathbf{b}(\mathbf{a}, z)$ attains its minimal value for P_Z-almost every $z\in Z$. In general, a minimax decision function is not necessarily invariant. We shall speak of a minimax invariant decision function defined above as a *Pitman-Girshick-Savage-Blackwell's decision function* (for shorten, *PGSB d.f.*).

The following theorem is also similar to Blackwell-Girshick's remark [4, p. 311].

Theorem 2.3. *The hypothesis (2.4) of Theorem 2.2 holds in each of the following cases:*

(1) *$w(\mathbf{a})$ is bounded.*

(2) *The decision space $\mathbf{A}$ is a topological space which has a sequence $\{C_n\}$ of compact subsets such that*

$$\bigcup_{n=1}^{\infty} C_n=\mathbf{A} \quad\text{and}\quad \lim_{n\to\infty}\inf_{\mathbf{a}\notin C_n} w(\mathbf{a})=\infty,$$

and moreover $w(\mathbf{a})$ is a continuous function of $\mathbf{a}\in\mathbf{A}$.

The proof is quite similar to that of Blackwell-Girshick.

2.4. Miscellaneous properties of PGSB d.f.

Theorem 2.4. *If $F^*<\infty$, then a PGSB d.f. φ^0 is a Bayes solution in the wide sense relative to a sequence $\{\xi_n\}$:*

$\xi_n(E)=\dfrac{\mu^r(K_n^*\frown L_E)}{\mu^r(K_n^*)}$ *for every Λ-measurable $E\subset\Omega$, where $L_E=\{\sigma:\sigma P\in E\}$ and K_n^* is a compact subset of G defined by (2.3), (2.4) and (2.10) for $J=J_n$ and $\varepsilon=\varepsilon_n\downarrow 0$.*

Proof. From (2.13) and (2.15) follows that,

$$\int_\Omega r(\sigma P, \varphi)\xi_n(d\sigma P)-\int_\Omega r(\sigma P, \varphi^0)\xi_n(d\sigma P)$$
$$=\frac{1}{\mu^r(K_n^*)}\left[\int_{K_n^*} r(\sigma P, \varphi)\mu^r(d\sigma)-\int_{K_n^*} r(\sigma P, \varphi^0)\mu^r(d\sigma)\right]$$
$$>(1-\varepsilon_n)(F^*-2\varepsilon_n)-F^*=-\varepsilon_n(F^*+2-2\varepsilon_n)$$

hold for every $\varphi\in\Phi$. Hence we have

$$\lim_{n\to\infty}\left\{\inf_{\varphi\in\Phi}\int_\Omega r(\sigma P, \varphi)\xi_n(d\sigma P)-\int_\Omega r(\sigma P, \varphi^0)\xi_n(d\sigma P)\right\}=0.$$

Corollary. If Ω is compact, a PGSB d.f. φ^0 is a Bayes solution relative to a normalized invariant distribution ξ on Ω, which is least favourable.

Proof. G is compact, since Ω is so. Hence, letting $K_n=G$, (2.13) yields

$$\int_\Omega r(\sigma P, \varphi)\xi(d\sigma P) \geqq F^* = \int_\Omega r(\sigma P, \varphi^0)\xi(d\sigma P)$$

for every $\varphi \in \Phi$, because G is unimodular and

$$\mu^r(\{\sigma : \sigma P \in E\}) = \xi(E).$$

Moreover we have

$$\inf_{\varphi\in\Phi} \int_\Omega r(\sigma P, \varphi)\pi(d\sigma P) \leqq \int_\Omega r(\sigma P, \varphi^0)\pi(d\sigma P)$$
$$= F^* = \inf_{\varphi\in\Phi} \int_\Omega r(\sigma P, \varphi)\xi(d\sigma P).$$

for any a priori distribution π, since $r(\sigma P, \varphi^0)$ has a constant value F^* for every $\sigma P \in \Omega$.

The admissibility of the PGSB d.f. φ^0 remains unsolved in the general case. Blackwell [2] has proved this in a special case. We have the other result, stated below, than Blackwell's.

Theorem 2.5. *If G is compact and if $W(\sigma P, \mathbf{a})$ is continuous on Ω uniformly in $\mathbf{a} \in \mathbf{A}$ and bounded on $\Omega \times \mathbf{A}$, then the PGSB d.f. is admissible.*

Proof. Since

$$| r(\sigma P, \varphi) - r(\tau P, \varphi) |$$
$$\leqq \left| \int_\Omega \sigma P(dx) \int_{\mathbf{A}} W(\sigma P, \mathbf{a})\varphi(d\mathbf{a} : x) - \int_\Omega \tau P(dx) \int_{\mathbf{A}} W(\sigma P, \mathbf{a})\varphi(d\mathbf{a} : x) \right|$$
$$+ \int_{\mathbf{A}} \tau P(dx) \int_{\mathbf{A}} \left| W(\sigma P, \mathbf{a}) - W(\tau P, \mathbf{a}) \right| \varphi(d\mathbf{a} : x),$$

we have the continuity of the function $r(\sigma P, \varphi)$ on Ω. Hence if $r(\sigma P, \varphi^0) \geqq r(\sigma P, \varphi)$ for all σP and if $r(\sigma P, \varphi^0) > r(\sigma P, \varphi)$ for some σP, then

$$F^* = r^*(\xi, \varphi^0) > r^*(\xi, \varphi),$$

because every open subset of Ω is of ξ-measure positive. This contradicts with the fact that φ^0 is a Bayes solution relative to ξ. (*cf.* Blyth [6, p. 28])

2.5. Some related problems.

a) *Coset estimates.* Let g be a closed subgroup of G. Suppose that $\mathbf{A}$ is a right coset space modulo g, and that the group $\tilde{G}$ of transformations of $\mathbf{A}$ consists of the transformations $\tilde{\rho} : g\sigma \to g\sigma\rho$. We have then

$$\mathbf{b}(g\sigma, z)=\int_G w(g\sigma\rho)\overline{P}(d\rho:z).$$

If $w(g\sigma)$ satisfies the condition (1) or (2) of Theorem 2.3, and if σ_z minimizes $\mathbf{b}(g\sigma, z)$, then the decision function φ^0:

(2.16) $$\varphi^0(g\sigma_z\rho:\rho x_z)=\varphi^0(g\sigma_z:x_z)=1$$

is a PGSB d.f. Under this decision function φ^0, the distribution of decision when P is true is

$$\begin{aligned}D(A:\varphi^0, P)&=\int_Z P_Z(dz)\int_G \varphi^0(A:\rho x_z)\overline{P}(d\rho:z)\\&=\int_Z P_Z(dz)\int_G \varphi^0(\tilde{\rho}^{-1}A:x_z)\overline{P}(d\rho:z)\\&=\int_Z \overline{P}(\{\rho:A\ni g\sigma_z\rho\}:z)P_Z(dz),\end{aligned}$$

where A is a Borel set of the right coset space G/g. Hence we have

Theorem 2.6. *If the decision space* $\mathbf{A}$ *is a right coset space* G/g, *then the distribution of decision under a non-randomized PGSB d.f. given as (2.16) when* σP *is true is independent of the parameters belonging to* $\bigcap_{z\in Z}\sigma_z^{-1}g\sigma_z$ *i.e.,*

$$D(A:\varphi^0, \sigma\tau P)=D(A:\varphi^0, \tau P)$$

for every $\sigma\in\bigcap_{z\in Z}\sigma_z^{-1}g\sigma_z$.

b) *The loss function for problems of estimation by intervals.* Suppose that the fiducial measure $\overline{P}$ for each z is absolutely continuous with respect to the left invariant Haar measure μ on G, and write

$$\overline{P}(L:z)=\int_L \overline{p}(\sigma:z)\mu(d\sigma).$$

Let the decision space $\mathbf{A}$ be the class of the subsets $\mathbf{a}$ of G transformed from the sets of the form $\{\sigma^{-1}:p(\sigma;z)\geqq u\}$ by elements $\tau\in G$. Put, for $\mathbf{a}\in\mathbf{A}$,

$$C(\mathbf{a})=\begin{cases}0 & \text{if } e\in\mathbf{a},\\ 1 & \text{if } e\notin\mathbf{a}.\end{cases}$$

Theorem 2.7. *If*

$$w(\mathbf{a})=\alpha C(\mathbf{a})+\beta\mu(\mathbf{a}^{-1}),\quad \alpha, \beta>0,$$

then the invariant decision function φ:

$$\varphi(\mathbf{a}_z:x_z)=1,\quad \mathbf{a}_z=\left\{\sigma^{-1}:\overline{p}(\sigma:z)\geqq\frac{\beta}{\alpha}\right\},$$

is a PGSB d.f.

Proof. $\mathbf{b}(\mathbf{a}, z)=\alpha\overline{P}(G-\mathbf{a}^{-1}:z)+\beta\mu(\mathbf{a}^{-1})=\alpha+\int_{\mathbf{a}^{-1}}\{\beta-\alpha\overline{p}(\sigma:z)\}\,\mu(d\sigma)$,

$\mathbf{b}_n(\mathbf{a}, z)=\alpha\overline{P}(J_n-\mathbf{a}^{-1}:z)+\beta\mu(\mathbf{a}^{-1})\overline{P}(J_n:z)=\alpha\overline{P}(J_n:z)+\beta\mu(\mathbf{a}^{-1}-J_n)\overline{P}(J_n:z)+\int_{\mathbf{a}^{-1}\frown J_n}\{\beta\overline{P}(J_n:z)-\alpha\overline{p}(\sigma:z)\}\mu(d\sigma)$. Denote $\mathbf{a}_{n,z}=\left\{\sigma^{-1}:\overline{p}(\sigma:z)\geqq\frac{\beta}{\alpha}\overline{P}(J_n:z)\right\}\frown J_n^{-1}$. Then obviously $\mathbf{a}_z$ and $\mathbf{a}_{n,z}$ minimize $\mathbf{b}(\mathbf{a}, z)$ and $\mathbf{b}_n(\mathbf{a}, z)$ respectively. From $\infty>\mu\left[\left\{\sigma:\overline{p}(\sigma:z)>\frac{\beta}{\alpha}\cdot\overline{P}(J_n:z)\right\}\right]\geqq\mu(\mathbf{a}_z^{-1})$, $\lim\overline{P}(J_n:z)=1$ and $\bigcup_{n=1}^{\infty}\mathbf{a}_{n,z}=\mathbf{a}_z$, it follows that there is for every $\varepsilon>0$ an integer n_0 such that $\mu(\mathbf{a}_{n,z}^{-1})\overline{P}(J_n:z)>\mu(\mathbf{a}_z^{-1})-\varepsilon$ and $\overline{P}(J_n-\mathbf{a}_{n,z}^{-1}:z)>\overline{P}(G-\mathbf{a}_z^{-1}:z)-\varepsilon$ hold for $n>n_0$. This shows that (2.4) holds and by Theorem 2.2 the PGSB d.f. exists. Hence by Corollary of Theorem 2.2 the proof is accomplished.

3. The relation between topologies of the sample space and of the distribution space.

3.1. *Homogeneous spaces with respect to a group.*

Suppose that i) the sample space X is a σ-compact, connected and locally compact metric space, where the metric in X is denoted by $((x, y))$;

ii) $\mathfrak{B}$ is the class of all Borel sets of X;

iii) G is a group of isometric transformations of X onto itself;

iv) X is homogeneous with respect to G;

v) if $\sigma_1 x, \sigma_2 x, \dots$ is a Cauchy sequence of sample points for every $x \in X$, then there exists an element $\sigma_0 \in G$ such that

$$\lim_{n\to\infty}\sigma_n x=\sigma_0 x \text{ for every } x \in X.$$

As seen in Theorems 1.1 and 1.3, G becomes a σ-compact and locally compact topological group with the base of neighbourhoods $u(k, \varepsilon)=\{\sigma: ((\sigma x, x))<\varepsilon$ for every $x \in k\}$ of the neutral element e, where k is a compact subset of X, and there is one and only one G-invariant measure $\mathfrak{m}$ on $(X, \mathfrak{B})$.

Further we suppose that

vi) the probability measure $P(B)$ is absolutely continuous with respect to the G-invariant measure $\mathfrak{m}$ stated above, and

$$P(B)=\int_B p(x)\mathfrak{m}(dx),\ B \in \mathfrak{B}.$$

From vi) every $\bar{\sigma}P$ $(\sigma\in G)$ is also absolutely continuous with respect to $\mathfrak{m}$ and we have

$$\bar{\sigma}P(B)=\int_B p(\sigma^{-1}x)\mathfrak{m}(dx),\ B \in \mathfrak{B}.$$

Lemma 3.1. *$V(\bar{\sigma}\bar{\tau}P, \bar{\tau}P)$ is a continuous function of σ for every fixed $\bar{\tau} \in \overline{G}$.*

Proof. Choose a compact subset k of X such that

$$\bar{\tau}P(X-k)<\varepsilon/3$$

for a given positive number ε. Such a set exists, for X is σ-compact and $\bar{\tau}P$ is finite measure. Let $v(x_0, 2\delta)$ be a neighbourhood of x_0 with radius 2δ, whose closure is compact. Since every transformation in G is isometric, the closure of every open set $v(x, 2\delta)$ is compact (from the condition iv)). Denote by $x_1, x_2, \ldots, x_n$ a finite δ-net covering k, and write $k'=\{x_1, x_2, \ldots, x_n\}$ and

$$k''=\text{the closure of } \bigcup_{j=1}^{n} v(x_j, 2\delta).$$

Then we have

$$u(k', \delta)k \subset k'',$$

because there is, for any $x \in k$ and any $\rho \in u(k', \delta)$, an $x_j \in k'$ such that

$$\begin{aligned}((\rho x, x_j)) &\leqq ((\rho x, \rho x_j)) + ((\rho x_j, x_j)) \\ &= ((x, x_j)) + ((\rho x_j, x_j)) < \delta+\delta=2\delta.\end{aligned}$$

From this fact, it holds for every $\rho \in u(k', \delta)$ that

$$(3.1) \qquad \begin{aligned}\bar{\rho}\bar{\tau}P(X-k'') &= \bar{\tau}P(X-\rho^{-1}k'') \\ &\leqq \bar{\tau}P(X-k)<\varepsilon/3.\end{aligned}$$

As seen in ix) of Theorem 1.4, we can choose a compact subset k^* of X and a positive number δ_1 such that

(3.2) [absolute variation of $(\bar{\sigma}\bar{\tau}P-\bar{\tau}P)$ on k''] $<\varepsilon/3$.

for every $\sigma \in u(k^*, \delta_1)$. Therefore it follows from (3.1) and (3.2) that

$$\begin{aligned}&V(\bar{\sigma}\bar{\tau}P, \bar{\tau}P) \\ &\leqq [\text{absolute variation of } (\bar{\sigma}\bar{\tau}-\bar{\tau}P) \text{ on } k''] + \bar{\sigma}\bar{\tau}P(X-k'') + \tau P(X-k'') \\ &< \frac{\varepsilon}{3} + \frac{2}{3}\varepsilon = \varepsilon .\end{aligned}$$

for every $\sigma \in u(k^*, \delta_1) \frown u(k', \delta)$. This completes the proof.

From Lemma 3.1. directly follows

Lemma 3.2. *The mapping $\sigma \rightarrow \bar{\sigma}P$ of G onto Ω is continuous. Hence Ω is a σ-compact metric space.*

Hence we have, by using the remark in the footnote 8),

Lemma 3.3. *There corresponds for every compact subset K of Ω and every real $\varepsilon>0$ a neighbourhood u of e of G such that*

$$\begin{aligned}\bar{u} &= \{\bar{\sigma} : \sigma \in u\} \\ &\subset U(K;\varepsilon) = \{\bar{\sigma} ; V(\bar{\sigma}\bar{\rho}P, \bar{\rho}P)<\varepsilon \text{ for every } \bar{\rho}P \in K\}.\end{aligned}$$

Remark 3.1. If we have known that Ω is locally compact, then the class of all $U(K, \varepsilon)$'s forms a complete system of neighbourhoods of $\bar{e}$, as seen in Theorem 1.1. In this case Lemma 3.3 asserts the continuity of the abgebraic homomorphism $\sigma \rightarrow \bar{\sigma}$, and hence, moreover, it asserts that $\sigma \to \bar{\sigma}$ is topologically a homomorphism. But we never know, up to this time, whether Ω is locally compact. Fortunately, however, we can prove the local compactness of Ω. We shall show this fact.

Lemma 3.4. *If $G_0 = \{e\}$, then the class of all $U(K, \varepsilon)$'s forms a complete system of neighbourhoods of e of G, which defines an equivalent topology to the original one of G.*

Proof. It is sufficient to show that the image $\bar{u}$ of every neighbourhood u of e contains at least one $U(K, \delta)$. We shall here show that there are a compact subset K and $\delta > 0$ for every $\varepsilon > 0$ and every $x_0 \in X$ such that

$$U(K, \delta) \subset \{\bar{\sigma} : ((\sigma x_0, x_0)) < \varepsilon\}.$$

This is sufficient for the proof (see the footnote 8)). Let us now suppose that there exists a sequence $\sigma_1, \sigma_2, \ldots$ of elements of G such that $((\sigma_n x_0, x_0)) > \varepsilon$ and $V(\overline{\sigma_n \rho} P, \bar{\rho} P) < \dfrac{1}{n}$ for $\bar{\rho} P \in K_n$ and $n = 1, 2, \ldots,$ where $K_1 \subset K_2 \subset \ldots,$ and $\lim K_n = \Omega$. Any subsequence of $\sigma_1, \sigma_2, \ldots$ has no limiting point in G. In fact, if a subsequence $\sigma_{n_1}, \sigma_{n_2}, \ldots$ tends to σ_0 in G, there is an integer j for every n such that $\overline{\sigma_0^{-1} \sigma_{n_i}} \in U\left(K_n, \dfrac{1}{n}\right)$ for every $i > j$ (Lemma 3.2). Hence we have $V(\overline{\sigma_0 \rho} P, \bar{\rho} P) \leqq V(\overline{\sigma_0 \rho} P, \overline{\sigma_{n_i} \rho} P) + V(\overline{\sigma_{n_i} \rho} P, \bar{\rho} P) < \dfrac{2}{n}$ for $\bar{\rho} P \in K_n$. From $\bigcup K_n = \Omega$ follows that $V(\overline{\sigma_0 \rho} P, \bar{\rho} P) = 0$ for every $\bar{\rho} P \in \Omega$. This means $\bar{\sigma}_0 = \bar{e}$, and hence $\sigma_0 = e$, since $G_0 = \{e\}$. This contradicts with $((\sigma_{n_i} x_0, x_0)) > \varepsilon$.

Let v be an open set $\{y : ((y, x_0)) < \varepsilon'\}$, whose closure is compact, and write $v_1 = \{y : ((y, x_0)) < \varepsilon'/3\}$. Since X is separable, there is an element $\bar{\sigma}_0 P$ of Ω for which

$$a = \bar{\sigma}_0 P(v_1) > 0.$$

Let $n_1, n_2, \ldots$ be a monotone increasing sequence of integers such that, for $j = 1, 2, \ldots,$

$$\frac{1}{n_j} < \frac{j}{j+1} a, \quad \bar{\sigma}_0 P \in K_{n_1},$$

and

$$\sigma_{n_j}^{-1} \in \bigcup_{i=0}^{j-1} u(\{\sigma_{n_i}^{-1} x_0\}, \varepsilon') \sigma_{n_i}^{-1},$$

where $\sigma_{n_0} = e$. Such a subsequence exists, since any subsequence of $\{\sigma_n\}$ has no limiting point and since the closure of $\sigma v = u(\{\sigma x_0\}, \varepsilon') \sigma x_0$ is compact for every $\sigma \in G$. For this subsequence $\{\sigma_{n_j}\}$ hold

$$\bar{\sigma}_0 P(\sigma_{n_j}^{-1} v_1) = \overline{\sigma_{n_j} \sigma_0} P(v_1) > \bar{\sigma}_0 P(v_1) - V(\overline{\sigma_{n_j} \sigma_0} P, \bar{\sigma}_0 P)$$

$$> \frac{a}{j+1},$$

and

$$\sigma_{n_j}^{-1} v_1 \frown [\bigcup_{i=0}^{j-1} \sigma_{n_i}^{-1} v_1] = \text{empty}.$$

This gives a contradiction:

$$1 = \bar{\sigma}_0 P(X) \geqq \bar{\sigma}_0 P(\bigcup_{j=0}^{\infty} \sigma_{n_j}^{-1} v_1) = \sum_{j=0}^{\infty} \bar{\sigma}_0 P(\sigma_{n_j}^{-1} v_1)$$
$$> a \sum_{j=0}^{\infty} \frac{1}{j+1} = \infty .$$

From Lemma 3.3 and 3.4 follows that, if $G_0 = \{e\}$, then $\overline{G}$ is a topological group and the algebraical isomorphism $\sigma \to \bar{\sigma}$ of G onto $\overline{G}$ is a homeomorphism.

Lemma 3.5. *The subgroup $H = \{\sigma : \bar{\sigma} P = P\}$ and hence G_0, being the meet of all $\rho H \rho^{-1}$, are closed in G.*

Proof. Let τ be an element of the closure of H. Then, by Lemma 3.3, there is a neighbourhood u of e for every given real $\varepsilon > 0$ such that $U(\{\bar{\tau} P\}, \varepsilon) \supset \bar{u}$. Since τu intersects with H, $\bar{\tau}\bar{u}$ and hence $\bar{\tau} U(\{\bar{\tau} P\}, \varepsilon)$ must intersect with $\overline{H}$. Let $\bar{\sigma}$ belong both of $\overline{H}$ and $\bar{\tau} U(\{\bar{\tau} P\}, \varepsilon)$. For such $\bar{\sigma}$, we have $\varepsilon > V(\bar{\sigma} P, \bar{\tau} P) = V(P, \bar{\tau} P)$. Since ε is arbitrary, $\bar{\tau} P = P$, *i.e.* $\tau \in H$. The closure of G_0 is evident.

By Lemma 3.5, Ω, being a coset space of $\overline{G}$, is σ-compact and locally compact, that is to say, Ω satisfies Assumption A, when $G_0 = \{e\}$.

Lemma 3.6. *Ω and $\overline{G}$ satisfy Assumption B, if $G_0 = \{e\}$.*

Proof. Suppose that $\overline{\sigma_1 \rho} P$, $\overline{\sigma_2 \rho} P$, ... is a Cauchy sequence in Ω for every fixed $\bar{\rho} P$. Since Ω is complete, there is a limiting point $\varphi(\rho P)$ of $\overline{\sigma_1 \rho} P$, $\overline{\sigma_2 \rho} P$, ... for every $\bar{\rho} P$. φ is evidently an isometric transformation of Ω onto itself. Denote by K the group of all isometric transformations of Ω. Since (Ω, K) satisfies Assumptions A and B, K becomes a complete topological group by Mibu's Theorems 1.1 and 1.2, and $\overline{G}$ is a closed subgroup of K, since G is complete. Hence there is an element $\bar{\sigma}_0$ of $\overline{G}$ such that $\varphi = \bar{\sigma}_0$. This shows that $\lim \overline{\sigma_n \rho} P = \overline{\sigma_0 \rho} P$ for every $\bar{\rho} \in \overline{G}$.

The subgroup I of the elements which remains a fixed sample point x_0 invariant is closed in G, and X is homeomorphic to the left coset space G/I under the mapping $\sigma x_0 \to \sigma I$, which can be proved similarly to Theorem 1.3. From this fact follows that the mapping $\sigma \to \sigma x_0$ of G onto X is open, and that $X_e = G_0 x_0$ is closed in X. Hence X_e is complete since X is complete.

Lemma 3.7. *$X_e = G_0 x_0$ is compact.*

Proof. Let $v(x_0, \varepsilon) = \{y : ((y, x_0)) < \varepsilon\}$ be a neighbourhood of x_0, whose closure is compact. If there exists an infinite sequence $x_0, x_1, x_2, \ldots$ in X_e, for which $((x_i, x_j)) > \varepsilon (i \neq j; i, j = 0, 1, 2, \ldots)$, then we have a disjoint sequence of open sets $v(x_i, \varepsilon/2)$, $i=0, 1, 2, \ldots$, each term of which satisfies

$$\bar{\sigma}P(v(x_i, \varepsilon/2)) = \bar{\sigma}P(\tau_i v(x_0, \varepsilon/2)) = \bar{\sigma}P(v(x_0, \varepsilon/2))$$

where $\bar{\sigma} \in \bar{G}$, $x_i = \tau_i x_0$ and $\tau_i \in G_0$. On the other hand, there is a $\sigma_0 \in G$ such that

$$\sigma_0 P(v(x_0, \varepsilon/2)) = P(\sigma_0^{-1} v(x_0, \varepsilon/2)) = a > 0,$$

since X is separable. This implies the contradiction:

$$1 = \bar{\sigma}_0 P(X) \geqq \bar{\sigma}_0 P(\bigcup_{i=1}^{\infty} v(x_i, \varepsilon/2)) = \sum_{i=1}^{\infty} a = \infty.$$

Thus we see that X_e has no infinite ε-net for any small $\varepsilon > 0$, that is to say, X_e is totally bounded. As seen above, X_e is complete, and hence compact.

Lemma 3.8. *Each $X_\sigma = \sigma G_0 x_0$ is homeomorphic to X_e under the mapping $\tau x_0 \to \sigma\tau x_0$ $(\tau \in G_0)$.*

Lemma 3.9. *The subgroup $G_0 I$ is closed in G.*

Proof. Let $\sigma_0 \notin G_0 I$. Then $\sigma_0 x_0 \notin X_e$. Hence there is a neighbourhood $v(\sigma_0 x_0, \varepsilon) = \{y : ((y, \sigma_0 x_0)) < \varepsilon\}$ disjoint with X_e. This shows that a neighbourhood $u(\{\sigma_0 x_0\}, \varepsilon)\sigma_0$ of σ_0 does not intersect with $G_0 I$.

Write

$$((X_\sigma, X_{\sigma'})) = \inf_{x \in X_\sigma, x' \in X_{\sigma'}} ((x, x')) = \inf_{\tau, \tau' \in G_0} ((\sigma\tau x_0, \sigma'\tau' x_0)) = \inf_{\tau \in G_0} ((x_0, \sigma^{-1}\sigma'\tau x_0)).$$

Obviously we have i) $((X_\sigma, X_{\sigma'})) \geqq 0$ with equality if and only if $\sigma^{-1}\sigma' \in G_0 I$, ii) $((X_\sigma, X_{\sigma'})) = ((X_{\sigma'}, X_\sigma))$, and iii) $((X_\sigma, X_{\sigma'})) + ((X_\sigma, X_{\sigma''})) \geqq ((X_{\sigma'}, X_{\sigma''}))$. We shall prove only iii). Given any real $\varepsilon > 0$, there correspond two elements τ and $\tau' \in G_0$ such that

$$((X_{\sigma'}, X_\sigma)) + \frac{\varepsilon}{2} > ((x_0, \sigma^{-1}\sigma'\tau x_0))$$

and

$$((X_{\sigma''}, X_\sigma)) + \frac{\varepsilon}{2} > ((x_0, \sigma^{-1}\sigma''\tau' x_0)).$$

Hence $((X_{\sigma'}, X_\sigma)) + ((X_\sigma, X_{\sigma''})) + \varepsilon > ((x_0, \sigma^{-1}\sigma'\tau x_0)) + ((x_0, \sigma^{-1}\sigma''\tau' x_0)) \geqq ((\sigma^{-1}\sigma'\tau x_0, \sigma^{-1}\sigma''\tau' x_0)) = ((\sigma'\tau x_0, \sigma''\tau' x_0)) \geqq ((X_{\sigma'}, X_{\sigma''}))$.

This metric $((X_\sigma, X_{\sigma'}))$ defines a topology of $Y = \{X_\sigma\}$, which is equivalent to the topology induced from the natural topology of the left coset space $G/(G_0 \cdot I)$ under the mapping $X_\sigma \to \sigma G_0 I$. On the other hand the factor group G/G_0 is regarded as a group of isometric transformations of Y.

Lemma 3.10. *If X and G satisfy the conditions,* i), iv), v) *on page 53, then Y and G/G_0 satisfy the same conditions.*

Proof. i, iv) are evident (see D and E on p. 60 of [20]). Let $\sigma_1, \sigma_2, \ldots$ be a sequence for which $\sigma_1 X_\sigma, \sigma_2 X_\sigma, \ldots$ forms a Cauchy sequence on Y for every X_σ. In the other word, $\sigma_1, \sigma_2, \ldots$ is a sequence, for which there corresponds a sequence $\tau_1, \tau_2, \ldots$ of elements of G_0 such that

$$((\sigma_n\tau_n\sigma x_0, \sigma_m\tau_m\sigma x_0)) \to 0 \text{ as } n, m \to \infty$$

for every $\sigma \in G$. By the condition v) assumed on X and G, there corresponds $\sigma_0 \in G$ such that

$$\lim_{n\to\infty} \sigma_n\tau_n\sigma x_0 = \sigma_0\sigma x_0 \text{ for every } \sigma \in G.$$

This means that

$$\lim_{n\to\infty} \sigma_n X_\sigma = \sigma_0 X_\sigma \text{ for every } X_\sigma \in Y.$$

By this Lemma, G/G_0 can be considered as a σ-compact and locally compact topological group. This topology of G/G_0 is equivalent to the natural topology of the factor group of the topological group G modulo G_0, because of the equality $\{\rho:((\rho X_\sigma, X_\sigma))<\varepsilon \text{ for } \sigma \in k \text{ (compact set)}\} = \{\rho:((\rho\sigma x_0, \sigma x_0))<\varepsilon \text{ for } \sigma \in k\}\cdot G_0$ and of the compactness of $\{\sigma x_0 : \sigma \in k\}$ and of $\{X_\sigma : \sigma \in k\}$ in the respective spaces X and Y.

From Lemma 3.4 follows that G/G_0, a group of isometric transformations of Y, is isomorphic to $\overline{G^Y}$, a group of isometric transformations of Ω^Y, as a topological group, if P_Y is absolutely continuous with respect to the invariant measure on $Y=\{X_\sigma\}$ (see (3.4) below). Hence the homomorphism $\sigma \to \bar{\sigma}_Y$ of G onto $\overline{G^Y}$ is open and continuous.

Define $\mathfrak{m}_Y(C)=\mathfrak{m}(\{\sigma : X_\sigma \in C\})$ for every Borel subset C of Y. This measure is a unique G/G_0-invariant measure on Y.

Since X_σ is a homogeneous space with respect to G_0, and since G_0 (or more precisely $G_0/(\bigcap_{\tau\in G_0} \tau\sigma I\sigma^{-1}\tau^{-1})$) can be regarded as a group of isometric transformations of the compact space X_σ which satisfies the conditions i)–iv) of page 53, there is one and only one G_0-invariant measure $\mathfrak{n}_\sigma$ of total measure 1 on each X_σ (see Lemmas 3.7 and 3.8). These measures fulfil evidently the relation

$$\mathfrak{n}_\sigma(B_\sigma)=\mathfrak{n}_e(\sigma^{-1}B_\sigma),\ B_\sigma=B\frown X_\sigma,\ B \in \mathfrak{B},$$

and hence it holds that

$$\mathfrak{m}(B\frown C^*)=\int_C \mathfrak{n}_\sigma(B_\sigma)\mathfrak{m}_Y(dX_\sigma), \tag{3.3}$$

where $B_\sigma = B\frown X_\sigma$ and $C^*=\{\sigma x_0 : X_\sigma \in C\}$. In fact, let

$$Q(B\frown C^*)=\int_C \mathfrak{n}_\sigma(B_\sigma)\mathfrak{m}_Y(dX_\sigma).$$

Then

$$\begin{aligned}Q(\rho(B\frown C^*))&=Q(\rho B\frown(\rho C)^*)\\&=\int_{\rho C}\mathfrak{n}_e(\sigma^{-1}(\rho B)_\sigma)\mathfrak{m}_Y(dX_\sigma)\\&=\int_{\rho C}\mathfrak{n}_e(\sigma^{-1}\rho B_{(\rho^{-1}\sigma)})\mathfrak{m}_Y(dX_\sigma)\\&=\int_C \mathfrak{n}_e(\sigma^{-1}B_\sigma)\mathfrak{m}_Y(dX_\sigma)=Q(B\frown C^*).\end{aligned}$$

From the uniqueness of the G-invariant measure on X, we have (3.3).

For an absolutely continuous measure P with respect to $\mathfrak{m}$ on X, we have

$$\begin{aligned}P(B\frown C^*)&=\int_{B\frown C^*}p(x)\mathfrak{m}(dx)\\&=\int_C \mathfrak{m}_Y(dX_\sigma)\int_{B_\sigma}p(x)\mathfrak{n}_\sigma(dx).\end{aligned}$$

Since $P(X)=1$, the integral

$$\int_{X_\sigma}p(x)\mathfrak{n}_\sigma(dx)$$

has a finite value for $\mathfrak{m}_Y$-almost every X_σ. Write

$$p_{\mathfrak{n}}(X_\sigma)=\int_{X_\sigma}p(x)\mathfrak{n}_\sigma(dx).$$

Then

$$P(B\frown C^*)=\int_C\frac{\int_{B_\sigma}p(x)\mathfrak{n}_\sigma(dx)}{\int_{X_\sigma}p(x)\mathfrak{n}_\sigma(dx)}p_{\mathfrak{n}}(X_\sigma)\mathfrak{m}_Y(dX_\sigma)$$

and

$$P_Y(C)=P(C^*)=\int_C p_{\mathfrak{n}}(X_\sigma)\mathfrak{m}_Y(dX_\sigma). \tag{3.4}$$

Hence we can define the conditional probability as

$$P(B:X_\sigma)=\frac{\int_{B_\sigma}p(x)\mathfrak{n}_\sigma(dx)}{\int_{X_\sigma}p(x)\mathfrak{n}_\sigma(dx)}.$$

From the definition of G_0 we have

$p(\tau x)=p(x)$ $\mathfrak{m}$-almost everywhere on X for $\tau\in G_0$.

Let $S_\sigma=\{x\,(\in X_\sigma):p(x)/p_{\mathfrak{n}}(X_\sigma)>1\}$, and $R_\sigma=\{x\,(\in X_\sigma):p(x)/p_{\mathfrak{n}}(X_\sigma)<1\}$. If there is a subset C of Y which is of positive $\mathfrak{m}_Y$-measure, and on each X_σ of which $\mathfrak{n}_\sigma(S_\sigma)>0$ and $\mathfrak{n}_\sigma(R_\sigma)=0$ hold, then $P(X)>1$. This is impossible. Hence $\mathfrak{n}_\sigma(S_\sigma)>0$ implies $\mathfrak{n}_\sigma(R_\sigma)>0$ for $\mathfrak{m}_Y$-almost all X_σ. However this is a contradiction, for there exists an element $\tau \in G_0$ such that $\mathfrak{n}_\sigma(\tau S_\sigma \frown R_\sigma) > 0$, from x) of Theorem 1.4.

Thus we have

$$p(x) = p_{\mathfrak{n}}(X_\sigma) \text{ for } x \in X_\sigma, \ \mathfrak{m}\text{-almost everywhere on } X,$$

and hence

$$P(B:X_\sigma) = \mathfrak{n}_\sigma(B_\sigma) \ \mathfrak{m}_Y\text{-almost everywhere on } Y.$$

From this follows

Lemma 3.11. *A G_0-invariant statistic exists and satisfies 3), 4) and 5) of Assumption C. And*

$$P(B \frown C^*) = \int_C \mathfrak{n}_\sigma(B_\sigma) p_{\mathfrak{n}}(X_\sigma) \mathfrak{m}_Y(dX_\sigma). \tag{3.5}$$

Therefore, from this lemma and b) of section 1.2, we can see that the mapping $\bar{\sigma}P \to \bar{\sigma}_Y P_Y$ is isometric and hence $\overline{G}_0^Y = \{\bar{e}\}$, and that $\overline{G}$ is isomorphic to $\overline{G}^Y$ as a topological group.

Lemma 3.12. *Suppose that g_0 is a locally compact group, and*

$$g_1 \supset g_2 \supset \ldots \supset g_n \supset \ldots$$

is a sequence of closed subgroups of g_0 such that g_{n+1} is a normal subgroup of $g_n (n=0, 1, 2, \ldots)$ and that

$$\bigcap_{n=1}^{\infty} g_n = \{e\}. \tag{3.6}$$

If factor groups $g_n/g_{n+1} (n=0, 1, \ldots,)$ are all compact, then g_0 is compact.

Proof. If g_0/g_n is compact, then g_0/g_{n+1} is so, since g_0/g_n is isomorphic to $(g_0/g_{n+1})/(g_n/g_{n+1})$ and since g_n/g_{n+1} is compact. By the mathematical induction every g_0/g_n is a compact group $(n=1, 2, \ldots)$. Denote by g^* the direct product of topological groups $g_0/g_1, g_0/g_2, \ldots$, and by g' be a closed subgroup of g^* consisting of all elements $(\sigma g_1, \sigma g_2, \ldots, \sigma g_n, \ldots)$: $\sigma \in g_0$. By Tichonoff's theorem g^* and hence g' are compact. By (3.6) the correspondence $\sigma \to (\sigma g_1, \sigma g_2, \ldots, \sigma g_n, \ldots)$ of g_0 onto g' is 1:1, and is evidently an open isomorphism. Hence the compactness of g_0 follows from that of g'.

Lemma 3.13. *The statement 2) of Assumption C is fulfiled.*

Proof. Let S be an everywhere dense countable subset of X, and write $J_n = \{\tau : \tau x_i = x_i, i=1, 2, \ldots, n\} \neq \{e\}$. Then the subset $S_n = \{x : J_n x = x\}$ of X is closed in X, since $y_i \in S_n$ and $y_i \to x''$ imply $x'' \in S_n$. Denote by x' one of the boundary points of S_n. Such a point exists because X is connected by the condition i). Let $v(x', \varepsilon)$ be a neighbourhood of x' whose closure is compact, and x_{n+1} a point of $(X\text{-}S_n) \frown v(x', \varepsilon) \frown S$. Since every element of J_n is a isometric transformation, $((\tau x_{n+1}, x')) = ((x_{n+1},$

$x'))<\varepsilon$ for every $\tau \in J_n$. Hence $J_n x_{n+1}$ is closed and contained in $v(x', \varepsilon)$ and hence compact. By Theorem 1.2 J_n/J_{n+1} is also compact.

By the repetition of the above procedure we have a sequence

$$J_1 \supset J_2 \supset \ldots \supset J_n \supset \ldots$$

of subgroups of G, where J_{n+1} is a normal subgroup of $J_n (n=1, 2, \ldots)$ and

$$\bigcap_{n=1}^{\infty} J_n = \{e\}. \qquad \text{(by the remark of the footnote 8).}$$

Hence the compactness of J_1 follows from Lemma 3.12.

Theorem 3.1. *Suppose that a sample space X and a transitive group G on X satisfy the conditions* i)–vi) *on page* 53. *Then Ω satisfies Assumptions A and B, and hence $\overline{G}$ can be topologized. Further the homomorphism $\sigma \to \bar{\sigma}$ of G is open and continuous, and the kernel of this homomorphism is $\overline{G}_0$. Assumptions C and D are fulfiled.*

Proof. From Lemmas 3.10, 3.11 and 3.4, $\overline{G^Y}$ can be topologized such that the isomorphism $\sigma G_0 \to \bar{\sigma}_Y$ of G/G_0 onto $\overline{G^Y}$ is topological. Since Ω^Y fulfils Assumption A (by Lemmas 3.10 and 3.5) and since $\bar{\sigma} P \to \bar{\sigma}_Y P_Y$ is an isometric mapping of Ω onto Ω^Y (by (3.5)), Ω also satisfies Assumption A. Hence $\overline{G}$ can be regarded as a topological group, where $U(K, \varepsilon)$ forms a complete system of neighbourhoods of $\bar{e}$, and the mapping $\bar{\sigma} \to \bar{\sigma}_Y$ is a topological isomorphism of $\overline{G}$ onto $\overline{G^Y}$ as seen in Section 1.2. Thus we can see easily that the homomorphism $\sigma \to \bar{\sigma}$ is open and continuous.

Now we shall prove that $(\Omega, \overline{G})$ satisfies Assumption B. By Lemma 3.6 and 3.10, (Ω^Y, G^Y) satisfies Assumption B. Suppose that $\overline{\sigma_1 \rho} P$, $\overline{\sigma_2 \rho} P$, $\ldots$ is a Cauchy sequence in Ω for every $\bar{\rho} P$. Since $\bar{\rho} P \to \bar{\rho}_Y P_Y$ is a isometric mapping, $\overline{(\sigma_1 \rho)}_Y P_Y$, $\overline{(\sigma_2 \rho)}_Y P_Y$, $\ldots$ is a Cauchy sequence in Ω^Y for evero $\bar{\rho}_Y P_Y$. Hence there is a $\bar{\sigma}_Y \in \overline{G^Y}$ such that

$$\lim_{n\to\infty} \overline{(\sigma_n \rho)}_Y P_Y = \bar{\sigma}_Y \overline{\rho_Y} P_Y, \ \bar{\rho}_Y P_Y \in \Omega^Y.$$

By applying the isometric mapping $\bar{\sigma}_Y P_Y \to \bar{\sigma} P$ again, we have

$$\lim_{n\to\infty} \overline{\sigma_n \rho} P = \overline{\sigma_0 \rho} P, \ \bar{\rho} P \in \Omega,$$

where σ_0 is an element such that $\overline{(\sigma_0)}_Y = \bar{\sigma}_Y$.

The last statement follows directly from Lemmas 3.11 and 3.13.

Remark 3.2. In this section we have assumed that the sample space X is a metric space, and that G is a group of isometric transformations of X. However a G-invariant metric can be defined in X, if X is a nuiform space satisfying the first axiom of coutability, where the neighbourhood $V_n(x)$ is symmetric and G-invariant, i.e.

a) $y \in V_n(x)$ implies $x \in V_n(y)$,

b) $\sigma V_n(x) = V_n(\sigma x)$ for all $\sigma \in G$,

c) $\bigcap_{n=1}^{\infty} V_n(x) = \{x\}$.

(see Bourbaki [5, Ch. IX, p. 22]).

3.2. *Reductions to a homogeneous space.*

Lemma 3.14. *Suppose that X is the direct product of two metric spaces Y and Z admitting transitive groups G and O of isometric transformations respectively. If X and the group $G' = G \times O$ satisfy the conditions* i)–v) *of the begining of the present section, then Y and G also satisfy these conditions. Further the topology of G induced by the metric of Y is equivalent to the relative topology of G as a subspace of the topological group G'.*

Proof. Since the conditions i)–iii) and v) are evidently satisfied by Y and G, we shall show only that iv) is fulfiled. If $\sigma_1 y, \sigma_2 y, \ldots$ is a Cauchy sequence of points for every $y \in Y$, then $(\sigma_1 y, z), (\sigma_2 y, z), \ldots$ is also a Cauchy sequence in X for every $y \in Y$ and $z \in Z$. Hence there exists an element of $\sigma_0 \in G$ such that

$$\lim_{n\to\infty} (\sigma_n y, z) = (\sigma_0 y, z) \quad \text{for every} \quad (y, z) \in X,$$

that is,

$$\lim_{n\to\infty} \sigma_n y = \sigma_0 y \quad \text{for every} \quad y \in Y.$$

The last statement of our Lemma follows directly from that $((\sigma\tau x, \tau x)) = ((\sigma x, x))$ holds for every $\sigma \in G$, $\tau \in O$ and $x \in X$.

Theorem 3.2. *Under the same assumptions as Lemma 3.14, Ω satisfies Assumption* A, *$\overline{G}$ becomes a topological group, and G is homomorphic to $\overline{G}$ as a topological group, if the initial probability measure P defined on X and the σ-field $\mathfrak{B}$ of all Borel subsets of X is absolutely continuous with respect to the G'-invariant measure on X.*

Proof. By Theorem 3.1. the mapping $\sigma \to \bar{\sigma}$ of G' onto $\overline{G'}$ is open and continuous, where $\overline{G'}$ is a group of transformations on the metric space $\Omega' = \{\bar{\sigma}P : \sigma \in G'\}$. Under this mapping G is mapped onto $\overline{G}$, and, since G is closed in G', $\overline{G}$ is closed in $\overline{G'}$. Hence Ω is closed in Ω', and satisfies Assumption A. From the relation $V(\sigma\tau\rho P, \tau\rho P) = V(\sigma\rho P, \rho P)$ for every $\tau \in O$ and $\sigma, \rho \in G$, it follows that the topology of $\overline{G}$ induced by the absolute variation metric V of Ω is equivalent to the relative topology of $\overline{G}$ as a subspace of $\overline{G'}$. Thus we can see that G is homomorphic to $\overline{G}$ under the mapping $\sigma \to \bar{\sigma}$ as a topological group.

Theorem 3.3. *Suppose that $(S, \mathfrak{S}, p)$ is a probability space, where $(S, \mathfrak{S})$ admits a group Γ of transformations satisfying the conditions* i)–v) *in the beginning of the present paper, and that p is absolutely continuous with respect to the Γ-invariant measure on S. Let $(X, \mathfrak{B}, P)$ be the direct product probability space of a finite number n of $(S, \mathfrak{S}, p)$, and G be the group of all transformations $(\sigma, \sigma, \ldots, \sigma):(s_1, s_2, \ldots, s_n) \to (\sigma s_1, \sigma s_2, \ldots, \sigma s_n)$. Then the distribution space $\Omega = \{\overline{(\sigma, \sigma, \ldots, \sigma)}P; \sigma \in \Gamma\}$ satisfies Assumption* A, *and hence $\overline{G}$ becomes a topological group (by Theorem 1.1), in each of the following cases:*

(a) *Γ is commutative,*

(b) *S is compact.*

Further, in such cases the mapping $\sigma \to (\sigma, \sigma, \sigma, \ldots, \sigma)$ of Γ onto $\overline{G}$ is open and continuous.

Proof. Denote by G' the direct product group $\Gamma \times \Gamma \times \ldots \times \Gamma$ of n groups $\Gamma:(\sigma_1, \sigma_2, \ldots, \sigma_n)\cdot(s_1, s_2, \ldots, s_n) = (\sigma_1 s_1, \sigma_2 s_2, \ldots, \sigma_n s_n)$. The space X and the group G' on X satisfy the conditions i)–v) in p. 53. Hence by Mibu's theorem (Theorems 1.1 and 1.2) G' can be topologized by the metric of X such that the operation of product is continuous. Evidently G is closed in G' and is isomorphic to Γ as a topological group.

Case (a). Denote by O the subgroup of G' consisting of the elements whose first coordinate σ_1 coincides with the neutral element e of Γ. Since Γ is commutative, and since the decomposition $(\sigma_1, \sigma_2, \ldots, \sigma_n) = (\sigma_1, \sigma_1, \ldots, \sigma_1)\cdot(e, \sigma_1^{-1}\sigma_2, \ldots, \sigma_1^{-1}\sigma_n)$ is unique, G' is the direct product of G and O. Moreover, fixing a point $(s^0, \ldots, s^0)$ of X, it follows from the commutativity of Γ that the mapping $(\sigma_1, \sigma_2, \ldots, \sigma_n) \to (\sigma_1 s^0, \ldots, \sigma_n s^0)$ of G' onto X is 1:1. Hence X may be considered as a direct product space $Y \times Z$: $Y = \{(s, s, \ldots, s): s \in S\}$ and $Z = \{(s^0, \sigma_2 s^0, \ldots, \sigma_n s^0): \sigma_i \in \Gamma, i = 2, \ldots n\}$, and G and O are transitive groups of transformations on Y and Z respectively. Therefore from Theorem 3.2 follows the conclusion.

Case (b). By Theorem 3.1 $\overline{G'}$, a group of transformations of $\Omega' = \{\bar{\sigma}P: \sigma \in G'\}$ induced by G', becomes a topological group and the mapping $(\sigma_1, \sigma_2, \ldots, \sigma_n) \to \overline{(\sigma_1, \sigma_2, \ldots, \sigma_n)}$ of G' onto $\overline{G'}$ is a homomorphism topologically. Hence $\sigma \to \overline{(\sigma, \sigma, \ldots, \sigma)}$ of Γ onto $\overline{G}$ as a subspace of G', is open and continuous. Since G is closed in G', $\overline{G}$ is closed in $\overline{G'}$. Hence Ω is closed in Ω', and satisfies Assumption A. Hence $\overline{G}$ becomes a topological group as a group of transformations of Ω, by the aid of the absolute variation metric of Ω. On the other hand, the identical mapping of the subgroup $\overline{G}$ of $\overline{G'}$ onto the group $\overline{G}$ of transformations of Ω is continuous, because every neighbourhood $U[\{\overline{(\sigma, \sigma, \ldots, \sigma)}P\}, \varepsilon]$, in $\overline{G}$, of $\bar{e}$ is the intersection of $\overline{G}$ and a neighbourhood, in $\overline{G'}$, of $\bar{e}$. From the compactness of S follows that of Γ, of G, and of the subgroup $\overline{G}$ of $\overline{G'}$. Therefore the mapping $\sigma \to (\sigma, \sigma, \ldots, \sigma)$ is open and continuous.

4. Some examples.

a) *The permutation parameter.* Let S be a set of k mutually exclusive events $s_1, s_2, \ldots, s_k$ with probability $p_1, p_2, \ldots, p_k$ ($\Sigma p_i = 1$) respectively, and denote such a probability space by (S, p). Consider the group Γ of the permutations σ of $1, 2, \ldots, k$. Every element σ of Γ introduces an another probability space $(S, \bar{\sigma}p)$ as follows:

$$\bar{\sigma}p(s_i) = p(s_{\sigma^{-1}(i)}) = p_{\sigma^{-1}(i)}; i = 1, 2, \ldots, k.$$

Suppose that $(X, \bar{\sigma}P)$ is a direct product probability space $(S \times S \times \times \ldots \times S, \bar{\sigma}p \times \bar{\sigma}p \times \cdots \times \bar{\sigma}p)$ of n same probability spaces $(S, \bar{\sigma}p)$, *i.e.* that X is a set of the sample points $x = (x_1, \ldots, x_n)$ in n independent trials, where the probability of occurrence of n_1 events of the first kind, n_2 of of the second kind, ... etc. ($\Sigma n_i = n$) in some order is give as

$$(4.1) \qquad p^{n_1}_{\sigma^{-1}(1)} p^{n_2}_{\sigma^{-1}(2)} \ldots p^{n_k}_{\sigma^{-1}(k)}. \text{[12)]}$$

Let G' be the group of the transformations $\sigma: (x_1, \ldots, x_n) \to (\sigma x_1, \sigma x_2, \ldots, \sigma x_n)$ of X for $\sigma \in \Gamma$, and G_0 the group of the permutations of the coordinates of $x = (x_1, x_2, \ldots, x_n)$. Evidently every element of G_0 leaves (4.1) invariant, and the group $G = G' \times G_0$ of transformations on X induces a distribution space $\Omega = \{\bar{\sigma}P : \sigma \in G\}$. Hence the G_0-invariant statistic $y(x) = (n_1, n_2, \ldots, n_k)$ for the above x is sufficient for Ω, and the distribution $\bar{\sigma}_Y P_Y$ on the range Y of $y(x)$ is given as

$$\bar{\sigma}_Y P_Y[(n_1, n_2, \ldots, n_k)] = \frac{n!}{n_1!\, n_2! \ldots n_k!} p^{n_1}_{\sigma^{-1}(1)} p^{n_2}_{\sigma^{-1}(2)} \ldots p^{n_k}_{\sigma^{-1}(k)},$$

when $\bar{\sigma}P$ is the true distribution of x.

If $(n_1, n_2, \ldots, n_k)$ is a permutation of a non-decending sequence $(m_1, m_2, \ldots, m_k)$, we put

$$z(n_1, n_2, \ldots, n_k) = (m_1, m_2, \ldots, m_k)$$

We can see easily that $z(y(x))$ is a G-invariant statistic, and that its distribution is

$$P_Z(m_1, m_2, \ldots, m_k) = \frac{n!}{m_1!\, m_2! \ldots m_k!} \sum p_1^{n_1} p_2^{n_2} \ldots p_k^{n_k}$$

irrespectively of the distribution of x, where the summation is over all different permutations $(n_1, n_2, \ldots, n_k)$ of $(m_1, m_2, \ldots, m_k)$. Hence the conditional probability of $(n_1, n_2, \ldots, n_k)$ for given $z(y) = (m_1, m_2, \ldots, m_k)$ when σP is true is

(12) For the simplicity we shall assume that (4.1) has a distinct value for a distinct set $(n_1, n_2, \ldots n_k)$.

$$\frac{p_{\tau^{-1}(1)}^{n_1} p_{\sigma^{-1}(2)}^{n_2} \dots p_{\sigma^{-1}(k)}^{n_k}}{\sum p_1^{n_1} p_2^{n_2} \dots p_k^{n_k}}.$$

Evidently $\overline{G}$ is a finite group, and hence compact. The fiducial measure $\overline{P}(\bar{\sigma}:z)$ is such that

$$\overline{P}(\bar{\sigma}:z)=\frac{p_1^{m_{\sigma(1)}} p_2^{m_{\sigma(2)}} \dots p_k^{m_{\sigma(k)}}}{\sum\limits_{\sigma\in\Gamma} p_1^{m_{\sigma(1)}} p_2^{m_{\sigma(2)}} \dots p_k^{m_{\sigma(k)}}}.$$

Suppose for example that $\mathbf{A}=\Omega=\{\bar{\sigma}P\}$ and $W(\bar{\sigma}P, \bar{\tau}P)=\sum\limits_{i=1}^{k}(p_{\sigma^{-1}(i)}-p_{\tau^{-1}(i)})^2$. Then we have

$$\begin{aligned}\mathbf{b}(\bar{\sigma}P, z)&=\sum_{\rho\in\Gamma}\sum_{i=1}^{k}(p_i-p_{\sigma^{-1}\rho^{-1}(i)})^2\overline{P}(\bar{\rho}:z)\\ &=\sum_{\rho\in\Gamma}\sum_{i=1}^{k}(p_i^2+p_{\sigma^{-1}\rho^{-1}(i)}^2-2p_i\,p_{\sigma^{-1}\rho^{-1}(i)})\,\overline{P}(\bar{\rho}:z)\\ &=2\sum_{\rho\in\Gamma}\sum_{i=1}^{k}(p_i^2-p_ip_{\sigma^{-1}\rho^{-1}(i)})\,\overline{P}(\bar{\rho}:z)\\ &=2[(\sum_{i=1}^{k}p_i^2)-(\sum_{i=1}^{k}p_i\sum_{\rho\in\Gamma}p_{\sigma^{-1}\rho^{-1}(i)}\,\overline{P}(\bar{\rho}:z))]\\ &=2[\sum_{i=1}^{h}p_i^2-b'(\sigma)],\ \text{(say).}\end{aligned}$$

Hence the PGSB d.f. is such that if the observed number $(n_1, n_2, \dots, n_k)$ is a result by permutation σ of $z=(m_1, m_2, \dots, m_k)$, we decide with probability 1 that $\bar{\sigma}\bar{\tau}_zP$ is the true distribution, where τ_z maximizes

$$b'(\tau)=\sum_{i=1}^{k}p_i\sum_{\rho\in\Gamma}p_{\tau^{-1}\rho^{-1}(i)}\,\overline{P}(\bar{\rho}:z)$$

For $k=2$,

$$b'(\sigma_0)\,M=p_1^{m_1}p_2^{m_2}(p_1^2+p_2^2)+2p_1^{m_2+1}p_2^{m_1+1}$$

for the identical permutation σ_0, and

$$b'(\sigma_1)\,M=2p_1^{m_1+1}p_2^{m_2+1}+(p_1^2+p_2^2)\,p_1^{m_2}p_2^{m_1}$$

for $\sigma_1=(12)$, where $M=p_1^{m_1}p_2^{m_2}+p_1^{m_2}p_2^{m_1}$. Hence if $p_1<p_2$, then

$$p_1^{m_1}p_2^{m_2}\geqq p_1^{m_2}p_2^{m_1}$$

and we have $b'(\sigma_0)\geqq b'(\sigma_1)$. This shows that the estimating process, which indicates σ_1P or P according as the number n_1 of occurrences of the event s_1 is greater or smaller than n_2 of s_2, is a PGSB d.f.

For $k=3$, suppose that $p_1<p_2<p_3$, and denote

$$a_0=p_1^2+p_2^2+p_3^2,\quad a_1=p_1^2+2p_2p_3,\quad a_2=p_2^2+2p_1p_3,$$
$$a_3=p_3^2+2p_1p_2,\quad a_4=p_1p_2+p_2p_3+p_3p_1,$$

and

$$P_0=p_1^{m_1}p_2^{m_2}p_3^{m_3},\quad P_1=p_1^{m_1}p_2^{m_3}p_3^{m_2},\quad P_2=p_1^{m_3}p_2^{m_2}p_3^{m_1},$$
$$P_3=p_1^{m_2}p_2^{m_1}p_3^{m_3},\quad P_4=p_1^{m_2}p_2^{m_3}p_3^{m_1},\quad P_5=p_1^{m_3}p_2^{m_1}p_3^{m_2}.$$

Then we have

$$(4.2)\qquad a_0>a_1>a_4>a_2,\quad a_0>a_3>a_4>a_2,$$
$$P_0>P_1>P_4>P_2,\quad P_1>P_5,$$
$$(4.3)\qquad P_0>P_3>P_5>P_2,\quad P_3>P_4.$$

Since the group G' consists of elements $\sigma_0=identical$, $\sigma_1=(12)$, $\sigma_2=(23)$, $\sigma_3=(31)$, $\sigma_4=(132)$ and $\sigma_5=(123)$, we have

$$b'(\sigma_0)=a_0P_0+a_1P_1+a_3P_3+a_4P_4+a_4P_5+a_2P_2,$$
$$b'(\sigma_1)=a_3P_0+a_4P_1+a_0P_3+a_2P_4+a_1P_5+a_4P_2,$$
$$b'(\sigma_2)=a_1P_0+a_0P_1+a_4P_3+a_3P_4+a_2P_5+a_4P_2,$$
$$b'(\sigma_3)=a_2P_0+a_4P_1+a_4P_3+a_1P_4+a_3P_5+a_0P_2,$$
$$b'(\sigma_4)=a_4P_0+a_3P_1+a_2P_3+a_0P_4+a_4P_5+a_1P_2,$$
$$b'(\sigma_5)=a_4P_0+a_2P_1+a_1P_3+a_4P_4+a_0P_5+a_3P_2,$$

Therefore, following from (4.2) and (4.3)

$$b'(\sigma_0)=\max_{0\leqq i\leqq 5} b'(\sigma_i).$$

This shows that the PGSB d.f. is a non-randomized estimation such that if the triplet (n_1, n_2, n_3) of the numbers of occurrences of the events s_1, s_2, s_3 is a permutation σ of (m_1, m_2, m_3), then the distribution $\bar{\sigma}P$ is accepted.

For $k\geqq 4$, calculations are rather troublesome.

b) *The discrete location parameter*. Let X be Euclidean n-space R^n, and P is a discrete distribution which distributes on $x_{ij}=y_i+r_{ij}x_0$ the probability s_ip_{ij}, where $y_i=(y_{i1}, y_{i2}, \ldots, y_{in})$ is a point in the hyperplane $\sum_{l=1}^{n} y_{il}=0$, $x_0=(1, 1, \ldots, 1)$, $s_i>0$, $\sum_{i=1}^{\infty} s_i=1$, $p_{ij}\geqq 0$, $\sum_{j=1}^{\infty} p_{ij}=1$, and $i, j=1, 2, \ldots$.

The group G of translations $h: x\rightarrow x+hx_0$ of X induces the distribution space $\Omega=\{\bar{h}P\}$:

$$\bar{h}P(x_{ij}+hx_0)=s_ip_{ij}.$$

However Ω is discrete and non-countable, in fact

$$V(\bar{h}P, P)\geqq \sup_{\substack{1\leqq i<\infty\\ 1\leqq j<\infty}}\ \inf_{\substack{1\leqq i'<\infty\\ 1\leqq j'<\infty}}\ \max\,[\,|s_ip_{ij}-s_{i'}p_{i'j'}|,\ s_ip_{ij}]>0,$$

whatever h may be. Hence Ω is not necessarily σ-compact, so that Mibu's Theorems 1.1–1.4 can not be directly applied for this problem.

Let $h_1, h_2, \ldots$ be a sequence of real numbers such that

1) $n_1h_1+n_2h_2+\ldots+n_kh_k=0$ implies $n_1=n_2=\ldots=n_k=0$ for every integer $k\geqq 1$,

2) for every r_{ij} there is one and only one linear expression

$$r_{ij}=\sum_{\alpha=1}^{k} n_\alpha h_\alpha$$

with integral coefficients $n_1, n_2, \ldots, n_k$.

Denote by G' the normal subgroup of G consisting of all linear combination $\sum n_\alpha h_\alpha$ of every finite number of $h_1, h_2, \ldots, h_k$ with integral coefficients $n_1, n_2, \ldots, n_k$. The distribution space Ω', induced from P by G', is a discrete but countable space, and hence is σ-compact. Evidently $\overline{G'}$ may be considered to coincide with G'. Let J_k be the subset of G' consisting of all $\sum_{\alpha=1}^{k} n_\alpha h_\alpha$ with $|n_\alpha|<k(\alpha=1, 2, \ldots, k)$ for a fixed positive integer k. Then we have

$$\bigcup_{k=0}^{\infty} J_k=G'$$

and J_k is a finite set. Hence there is an integer $k>0$ for every finite subset R such that $J_k\supset R$. It is obvious that G' is an A-group. Hence, by Theorem 2.2, we have

$$\sup_{\bar{h}P\in\Omega} r(\bar{h}P, \varphi) \geqq \sup_{\bar{h}P\in\Omega'} r(\bar{h}P, \varphi) \geqq \sum_{i=1}^{\infty} F_i \quad \text{for every d.f. } \varphi$$

if there is a finite subset R of G' for every $\varepsilon>0$ and every integer i $(0<i<\infty)$ such that

$$\sum_{r_{ij}\in R} w(h+r_{ij})p_{ij} > F_i-\varepsilon \quad \text{for every } h \in G',$$

where $F_i=\inf_{h\in G'} \sum_{j=1}^{\infty} w(h+r_{ij})p_{ij}$. This is Theorem 11.3.1. of Blackwell-Girshick's Book [4].

The admissibility of this d.f. has shown by Blackwell [2] under a preferably weak condition.

c) *The continuous location parameter.* Let $X=R^n(n\geqq 2)$, $\mathfrak{B}$ the class of all Borel sets of R^n, and P a probability measure with a density function $f(x)=f(x_1, x_2, \ldots, x_n)$:

$$P(B)=\int_B f(x_1, x_2, \ldots, x_n)\, dx_1 dx_2 \ldots dx_n, \quad B \in \mathfrak{B},$$

where the integration is the usual Lebesgue's one.

Suppose that G is a group of transformations $(x_1, x_2, \ldots, x_n)\to(x_1+\sigma, x_2+\sigma, \ldots, x_n+\sigma)$, where σ is a real number. This group G induces a distribution space $\Omega=\{\bar{\sigma}P: -\infty<\sigma<\infty\}$ from P as follows:

$$\begin{aligned}\bar{\sigma}P(B)&=\int_{B-\sigma} f(x_1, x_2, \ldots, x_n)\, dx_1 dx_2 \ldots dx_n \\ &=\int_B f(x_1-\sigma, x_2-\sigma, \ldots, x_n-\sigma)\, dx_1 dx_2 \ldots dx_n.\end{aligned}$$

Such a parameter σ is called a *location parameter*.

By Theorem 3.2 the mapping $\sigma \to \bar{\sigma}$ of G onto the parameter group $\bar{G}$ is isomorphic topologically, that is, $\bar{G}$ is also considered as the additive group of the real numbers. Hence $\bar{G}$ is an A-group.

The fiducial measure $\bar{P}(L:z)$ for $z=(0, z_2, \ldots, z_n)$ is

$$\bar{P}(L:z)=\int_L f(\sigma, z_2+\sigma, \ldots, z_n+\sigma)d\sigma / h(z)dz.$$

Suppose that $\mathbf{A}=\Omega$ and $W(\bar{\sigma}P, \bar{\tau}P)=w(\tau-\sigma)$, and write

$$\mathbf{b}_m(\tau, z)=\int_{-m}^{+m} w(\sigma+\tau)f(\sigma, z_2+\sigma, \ldots, z_n+\sigma)d\sigma,$$

$$\mathbf{b}(\tau, z)=\int_{-\infty}^{+\infty} w(\sigma+\tau)f(\sigma, z_2+\sigma, \ldots, z_n+\sigma)d\sigma,$$

where $z=(0, z_2, \ldots, z_n)$. If there is, for every z and every positive ε, an integer m such that

$$\mathbf{b}_m(\tau, z) > \inf_{-\infty<\tau<+\infty} \mathbf{b}(\tau, z)-\varepsilon \quad \text{for every real } \tau,$$

and if τ_z minimizes $\mathbf{b}(\tau, z)$, then the PGSB d.f. exists and is the procedure in which, when $x=(x_1, x_2, \ldots, x_n)$ is observed, $(\overline{x_1+\tau_z})P$ is accepted, where $z=(0, x_2-x_1, \ldots, x_n-x_1)$.

i) If $w(\sigma)=\sigma^2$ and if $\int_{-\infty}^{\infty}(\sigma+\tau)^2 f(\sigma, z_2+\sigma, \ldots, z_n+\sigma)\ d\sigma<\infty$ for all τ, then we have

$$\tau_z=-\int_{-\infty}^{+\infty}\sigma f(\sigma, z_2+\sigma, \ldots, z_n+\sigma)d\sigma, \ z=(0, z_2, \ldots, z_n).$$

The minimax risk of this decision function is given as

$$\int_{-\infty}^{++}\cdots\int_{-\infty}^{+\infty} x_1^2 f(x_1, x_2+x_1, \ldots, x_n+x_1)\,dx_1 dx_2 \ldots dx_n$$
$$-\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\tau_z^2 h(z)\,dz_2 \ldots dz_n.$$

ii) If $w(\tau)=|\tau|$, then τ_z is such that

$$\int_{-\infty}^{\tau_z} f(\sigma, z_2+\sigma, \ldots, z_n+\sigma)\,d\sigma=\tfrac{1}{2},\ z=(0, z_2, \ldots, z_n)$$

The above two estimates are the ones obtained by Pitman [18].

iii) If $f(\sigma, z_2+\sigma, \ldots, z_n+\sigma)$ is a monotone decreasing function $g_z(t)$ of $t=|\sigma-\sigma_z|$ for every fixed $z=(0, z_2, \ldots, z_n)$ and $w(\tau)$ is also a monotone increasing function of $|\tau|$, then the observed value $(x_1, x_2, \ldots, x_n)$ indicates the estimated value $x_1-\sigma_z$ for $z=(0, x_2-x_1, \ldots, x_n-x_1)$. This is a maximal likelihood estimate as well as a PGSB estimate.

iv) Suppose that $f(x_1, x_2, \ldots, x_n)$ satisfies the same conditions as the case iii), $\mathbf{A}$ is the class of all interval $I_{ab}=(a, b): -\infty<a<b<+\infty$, and that $w(I_{ab})=|b-a|$ if $a<0<b$; $=1+|b-a|$ if otherwise. For such

problems, the observed value $(x_1, x_2, \ldots, x_n)$ indicates the estimation by the interval $(\sigma_z-\sigma_1+x_1, \sigma_z+\sigma_1+x_1)$ for $z=(0, x_2-x_1, \ldots, x_n-x_1)$, where $g_z(t)\geqq 1$ for $|t|\leqq\sigma_1$ and $g_z(t)<1$ for $|t|>\sigma_1$. This procedure is minimax by Theorem 2.7.

d) *The location-scale parameter.* Let X, $\mathfrak{B}$ and P are the same as defined in c). Suppose that G is a group of transformations $\sigma(x_1, x_2, \ldots, x_n)=(\sigma x_1, \sigma x_2, \ldots, \sigma x_n)$, where $\sigma x_i=a_\sigma x_i+b_\sigma$, $+\infty>a_\sigma>0$, $+\infty>b_\sigma>-\infty$. ($a_{\sigma\tau}=a_\sigma a_\tau$ and $b_{\sigma\tau}=a_\sigma b_\tau+b_\sigma$). Then

$$\bar{\sigma}P(B)=P(\sigma^{-1}B)=\int_{\sigma^{-1}B} f(x_1, x_2, \ldots, x_n)\, dx_1 dx_2 \ldots dx_n$$
$$=\frac{1}{a_\sigma^n}\int_B f\left(\frac{x_1-b_\sigma}{a_\sigma}, \frac{x_2-b_\sigma}{a_\sigma}, \ldots, \frac{x_n-b_\sigma}{a_\sigma}\right)dx_1\, dx_2 \ldots dx_n.$$

We shall consider as Ω the whole of $\bar{\sigma}P$'s. In this case the pair (a_σ, b_σ) is called a *location-scale parameter.*

Write, for every point $x=(x_1, x_2, \ldots, x_n)$,

$$\bar{x}=\bar{x}(x)=\frac{\sum x_i}{n},\quad s=s(x)=\sqrt{\frac{\sum(x_i-\bar{x})^2}{n}}=\sqrt{\frac{\sum x_i^2}{n}-\bar{x}^2}$$

$$\text{and } z=z(x)=\left(\frac{x_1-\bar{x}}{s}, \frac{x_2-\bar{x}}{s}, \ldots, \frac{x_n-\bar{x}}{s}\right).$$

Denote by X_0 the line $x_1=x_2=\ldots=x_n$ and by x_0 a point $(1, 1, \ldots, 1)$. Since every $x\notin X_0$ is expressed uniquely as

$$x=sx+\bar{x}x_0,$$

and since the G-invariant statistic $z(x)$ has its values on the unit sphere Z of the hyperplane $z_1+z_2+\ldots+z_n=0$, we can write

$$f(x)=p(\bar{x}(x),\ s(x),\ z(x))$$

and

$$dx_1dx_2\ldots dx_n=s^{n-2}\,d\bar{x}dsdz,$$

where dz is an element of area on the unit $(n-2)$-dimensional sphere. Denoting by B' the image of $B(\in\mathfrak{B})$ under the mapping $x\to(\bar{x}(x), s(x), z(x))$, we can easily see that

$$\bar{\sigma}P(B)=\frac{1}{a_\sigma^n}\int_{B'} p\left(\frac{\bar{x}-b_\sigma}{a}, \frac{s}{a_\sigma}, z\right) s^{n-2}d\bar{x}ds\, dz.$$

Hence, by writing

$$h(z)=\int_0^\infty\left\{\int_{-\infty}^{+\infty} p(\bar{x}, s, z)d\bar{x}\right\}s^{n-2}\,ds,$$

the distribution on Z induced by the statistic $z(x)$ is

$$\bar{\sigma}P_Z(D)=P_Z(D)=\int_D h(z)dz$$

and the conditional distribution for given z is

$$\bar{\sigma}P(B:z)=\frac{1}{a_\sigma^n}\int_{B'\frown X'z}\frac{p\left(\frac{\bar{x}-b_\sigma}{a_\sigma},\frac{s}{a_\sigma},z\right)}{h(z)}s^{n-2}\,d\bar{x}ds,$$

where D is a Borel subset of Z and $X_z'=\{(\bar{x},s,z):-\infty<\bar{x}<+\infty,\ 0<s<+\infty\}$. Hence the fiducial measure for z is such that

$$\bar{P}(L:z)=\frac{1}{h(z)}\int_L p(\bar{x},s,z)s^{n-2}d\bar{x}ds.$$

Now consider the group O of the rotations around the fixed axis X_0. Since Z is a homogeneous space with respect to O, the algebraically homomorphic mapping $\sigma\to\bar{\sigma}$ of G onto $\bar{G}$ is open and continuous from Theorem 3.2. However no element ρ of G remains each $\bar{\sigma}P$ invariant, if ρ is not the netural element, by Lemma 3.7. Hence $\sigma\to\bar{\sigma}$ is an isomorphism as a mapping of topological groups.

Denote by Ω' the distribution space induced from P by the direct product $G'=G\times O$, and by G_0 the subgroup of G' remaining every distribution of Ω' invariant. Since the intersection of G_0 and G contains only the netural element as seen just now, G_0 must be a normal subgroup of O. But O has no closed normal subgroup except for O itself and $\{e\}$. Hence $G_0=O$ or $G_0=\{e\}$.

If $G_0=O$, then Ω' and $\bar{G}'$ coincide with Ω and $\bar{G}$, respectively. And $(\bar{x}(x), s(x))$ becomes a sufficient statistic for Ω.

i) Suppose that $\mathbf{A}=\bar{G}$, and that the loss function $W(P,\bar{\sigma})=w(a_\sigma,b_\sigma)$ is either a bounded function on the upper half plane $\bar{G}$ or a continuous function satisfying

$$\lim_{a\to 0}w(a,b)=\lim_{a\to\infty}w(a,b)=\infty\quad\text{for any fixed } b,$$

and

$$\lim_{b\to\pm\infty}w(a,b)=\infty\quad\text{for any fixed } a>0.$$

Let (a_z,b_z) be a point of $\bar{G}$ which minimizes

$$\mathbf{b}[(a,b),z]=\frac{1}{h(z)}\int_0^\infty\left[\int_{-\infty}^\infty w(as,sb+\bar{x})p(\bar{x},s,z)d\bar{x}\right]s^{n-2}ds.$$

Then the decision function φ^0:

$$\varphi^0[(sa_z,sb_z+\bar{x}):sz+\bar{x}x_0]=1\ \text{for every } s>0 \text{ and } \bar{x}$$

is a PGSB d.f.

ii) Let g_b is a normal subgroup $\{(1,b);-\infty<b<\infty\}$. If $\mathbf{A}=G/g_b$, $W(P,g_b(a,0))=w(a)$ is either a bounded function or a continuous function such that $\lim_{a\to 0}w(a)=\lim_{a\to\infty}w(a)=\infty$, then the invariant d.f. φ^0:

$$\varphi^0[(a_zs,0)\cdot g_b:sz+x_0]=1$$

is a PGSB d.f., where a_z minimizes

$$\mathbf{b}(a, z)=\frac{1}{h(z)}\int_0^{+\infty} w(as)\left[\int_{-\infty}^{+\infty} p(\bar{x}, s, z)d\bar{x}\right] s^{n-2}\, ds.$$

The distribution of decision under φ^0 when $\bar{\sigma}P$ is true is independent of b_σ. (See Theorem 2.6)

iii) Let g_a is a subgroup $\{(a, 0): 0<a<\infty\}$. If **A** is a right coset space of G modulo g_a, and if $W(P, g_a\cdot(1, b))=w(b)$ is either a bounded function or a continuous function such that $\lim_{b\to\pm\infty} w(b)=\infty$, then φ^0:

$$\varphi^0\left(g_a\cdot\left(\frac{\bar{x}+b_z}{s}\right); sz+\bar{x}x_0\right)=1,$$

where b_z minimizes

$$\mathbf{b}(b, z)=\int_{-\infty}^{+\infty}\int_0^{+\infty} w\left(\frac{\bar{x}+b}{s}\right)p(\bar{x}, s, z)\,s^{n-2}\,ds\,d\bar{x}.$$

The distribution of decision under this φ^0 when $\bar{\sigma}P$ is true depends only on $\dfrac{b_\sigma+b_0}{a_\sigma}$, if $b_z=b_0$ for every $z\in Z$. (See Theorem 2.6)

If P is a normal distribution, such a decision function is of Student type.

e) *The rotation parameter*. Given n independent and identically distributed random points $s_1, s_2, \ldots, s_n$ on the unit sphere S in R^3, whose distributions are absolutely continuous with respect to the element of area on S and has a density function $f(s)$, we consider the space X of such samples $x=(s_1, s_2, \ldots s_n)$ as a direct product space $S\times S\times\ldots\times S$ of n spheres. Denote by $\sigma, \tau, \ldots$ the rotations on S. The group G of transformations on X, whose elements are $(s_1, s_2, \ldots, s_n)\to(\sigma s_1, \sigma s_2, \ldots, \sigma s_n)$, induces a distribution space $\Omega=\{\bar{\sigma}P\}$:

$$\bar{\sigma}P(B)=P(\sigma^{-1}B)=\int_B \prod_{i=1}^n f(\sigma^{-1}s_i)\,ds_1\ldots ds_n.$$

By $M(x)$ we denote the $n\times n$ matrix $(m_{ij}(x))$ whose i, j element $m_{ij}(x)$ is the smaller angle between two segments joining the origin with s_i and s_j respectively. This statistic $M(x)$ is G-invariant, and its distribution does not depend on σ.

Suppose that $n\geq 3$. By Theorom 3.3 G is isomorphic to $\overline{G}$. Thus $\overline{G}$ is compact and hence A-group. Therefore for this problem the PGSB d.f. exists if the loss function is continuous.

f) *The general linear transformation parameter*. Let X be the set of all $n\times k$ matrices $x=(x_{ij})$ $(n>k)$, and P be a k-variate normal distribution:

$$P(dx)=\frac{1}{(2\pi)^{\frac{1}{2}nk}}\,e^{-\frac{1}{2}tr(x'x)}\prod_{i,j} dx_{ij}.^{13)}$$

(13) x' is the transposed matrix of x.

For this example, A.T. James' paper [8] may be referred. Since every regular $k\times k$ matrix $\sigma=(\sigma_{ij})$ defines a right transformation $x\to x\sigma$ of X onto itself, the group G consisting of all of these transformations induces the distribution space $\Omega=\{\bar{\sigma}P\}$:

$$\bar{\sigma}P(dx)=\frac{|\Sigma|^{-\frac{n}{2}}}{(2\pi)^{\frac{1}{2}nk}}e^{-\frac{1}{2}tr(\Sigma^{-1}x'x)}\prod_{i,j}dx_{ij}.$$

where $\Sigma=\sigma'\sigma$. That is to say, Ω consists of the sample distributions of the normal k-variates $x_{i1},\dots,x_{ik}$ with means zero.

X is not a homogeneous space with respect to G. However, if we introduce on this space X the group O of the left transformations $x\to\theta x$ by $n\times n$ orthogonal matrices θ, then X becomes a homogeneous space with respect to the direct product group $G'=G\times O$. Since X satisfies the conditions a), b) and c) of Remark 3.2, and iv), v) of the beginning of Section 3, and since the G'-invariant measure on X is

$$\frac{\prod_{i,j}dx_{ij}}{|x'x|^{\frac{n}{2}}},$$

the mapping $\sigma\to\bar{\sigma}$ of G onto $\bar{G}$ is a homomorphism, by Theorem 3.2. Evidently there is no right transformation (if not neutral) on X remaining every quadratic form $tr(\Sigma^{-1}x'x)$ invariant. This shows that G_0 consists only of the neutral element, and that $\sigma\to\bar{\sigma}$ is an isomorphism topologically.

Thus we see that $\bar{G}$ isomorphic to the general linear group G of the regular $k\times k$ matrices. But we cannot yet to prove that G is an A-group. Hence we never know whether or not our method of constructing the invariant minimax decision function is valid for this problem.

References

[1] Berger, A.: Remark on separable spaces of probability measures, *Annals of Math. Stat.* **22** (1951).

[2] Blackwell, D.: On the translation parameter problem for discrete variables, *Annals of Math. Stat.* **22** (1951).

[3] Blackwell, D.: Comparison of experiments, *Proc. Sec. Berkley Sym.* 1951.

[4] Blackwell, D. and Girshick, M. A.: *Theory of games and statistical decisions*, New York, 1954.

[5] Bourbaki, N.: *Topologie génerale*, Paris, (2nd. ed.) 1951.

[6] Blyth, C.R.: On minimax statistical decision procedures and their admissibility, *Annals of Math. Stat.* **22** (1951).

[7] de Finetti, B.: Recent suggestions for the reconcillation of theories of probability, *Proc. Sec. Berkley Sym.* 1951.

[8] Fisher, R. A.: Two new properties of mathematical likelihood, *Proc. Roy. Soc.*, ser. **A. 144** (1934).

[9] Girshick, M. A. and Savage, L. J.: Bayes and minimax estimates for quadratic loss functions, *Proc. Sec. Berkley Sym.* 1951.

[10] Halmos, P. R.: *Measure theory*, New York, 1950.

[11] James, A. T.: Normal multivariate analysis and the orthogonal group, *Annals of Math. Stat.* **25** (1954).

[12] Karlin, S.: The theory of infinite games, *Annals of Math.* **58** (1953).

[13] Kudō, H.: On a formulation of classical problems of statistics, *Nat. Sci. Rep. of Ochanomizu Univ.* **1** (1951).

[14] Kudō, H.: On the transformations of the population and invariant estimates, *Kōkyūroku of Inst. Stat. Math.* **5** (1949) (in Japanese, mimeographed).

[15] Kudō, H.: Experiments and their informations, *Theory of time series and informations and its applications*. Tokyo, 1953 (in Japanese).

[16] Lehmann, E. L.: Some principles of the testing hypotheses. *Annals of Math. Stat.* **21** (1950).

[17] Mibu, Y.: Generalization of the theory of Haar measure. Measures on uniform space which are invariant under a given group of homeomorphisms. *Sūgaku* **1** (1947) (in Japanese).

[18] Pitman, E. J. G.: The estimation of the location and scale parameters of a continuous population of any given form, *Biometrika* **30** (1938).

[19] Pitman, E. J. G.: Test of hypotheses concerning location and scale parameters. *Biometrika* **31** (1939).

[20] Pontrjagin, L.: *Topological groups*, Princeton, 1939.

[21] Robins, H.: Mixture of distributions, *Annals of Math. Stat.* **19** (1948).

[22] Wald, A.: *Statistical decision functions*, New York, 1950.

[23] Wolfowitz, J.: The power of the classical tests associated with the normal distribution, *Annals of Math. Stat.* **20** (1949).

[24] Peisakoff, M. P.: Transformation parameters (Abstract), *Annals of Math. Stat.* **22** (1951).

(*Received August* 17, 1955)