Z. Wahrscheinlichkeitstheorie verw. Geb. 10, 70—78 (1968)

Statistically Inspired Conditions on the Group Structure of Invariant Experiments and their Relationships with Other Conditions on Locally Compact Topological Groups *

M. Stone and R. von Randow

Received November 11, 1967

Summary. In this paper are considered some conditions on locally compact topological groups that have arisen from the statistical analysis of group-invariant experiments and from probability theory. A catalogue is compiled of these conditions and of similar conditions that arise in the area of invariant means on locally compact topological groups. Known facts about the implication diagram of the conditions are described and some new implications are established. Counter-examples are discussed. An attempt is then made to throw some new light on the statistical relevance of the conditions in test theory, information theory and decision theory.

§ 1. Introduction

The last twenty years have witnessed a rapid incursion of group theory into probability and mathematical statistics. A good background account is provided by Lehmann [18], Hannan [11] and Grenander [10].

This paper will be concerned with groups that are associated with invariant experiments as follows.

As Lehmann, [18] p. 214, shows, the four suppositions

(i) data $x \in X$ has probability distribution P_θ, $\theta \in \Theta$,

(ii) there is a set $G = \{g\}$ of $1-1$ transformations $g\colon X \to X$ such that, if x has distribution P_θ, then gx has distribution $P_{\bar\theta}$ for some $\bar\theta = \bar\theta(\theta)$,

(iii) $\{\bar\theta(\theta) \mid \theta \in \Theta\} = \Theta$,

(iv) $P_{\theta_1} \neq P_{\theta_2}$, $\theta_1 \neq \theta_2$

imply that G is a group. For obvious reasons, the experiment may be said to be invariant under each $g \in G$.

The further supposition

(v) G is a locally compact topological group

will be made in this paper. It is difficult to imagine that there can be any problem of statistical interest in which (v) does not hold.

Statistical research on such invariant experiments may be divided into two classes. On the one hand, we have work in which the principal results apply indiscriminately to all locally compact topological groups, for example, Brillinger [1], Fraser [6], Hora and Buehler [13]. On the other hand, there is the class of results that apply to only a subclass of the family of locally compact

* Research partly supported by the Office of Naval Research while the first author was at the University of Wisconsin.

topological groups; for example, HUNT and STEIN [18] p. 336, PEISAKOFF [22], KUDO [17], KIEFER [16], GRENANDER [10] p. 104, and STONE [27].

The conditions on G arising in this latter area from the object of investigation of this paper. The question of whether similar conditions have been considered in other applications of group theory has by now a strongly affirmative answer. As early as 1929, VON NEUMANN [20] showed that not all locally compact topological groups possessed a left invariant mean. Later group theoretical studies by CALDERON [2], DAY [3], DIEUDONNÉ [4, 5], FØLNER [7], GODEMENT [8], GREENLEAF [9], HULANICKI [14], KESTEN [15], NAMIOKA [19], REITER [23, 24, 25] and STEGEMAN [26] have considerably clarified the connexions. A good account of much of this is to be found in HEWITT and ROSS [12]. However, as far as we know, it is still an open question whether or not all of the apparently different conditions are equivalent.

The purposes of this paper are to review the whole area, to establish new links, to show by a counter-example that one of the two potentially weakest conditions we present is not vacuous (the other is known to be non-vacuous by indirect argument) and then briefly to review and extend the statistical implications of some of the conditions in an attempt to achieve an appreciation of their importance.

Sections 2 and 3 are entirely non-statistical. Here we have not hesitated to construct small variants of some conditions when, by so doing, their relationships become more comprehensible. Section 4 is statistical and contains some speculative but, it is hoped, suggestive material.

§ 2. The Conditions and their Connexions

Some necessary definitions are:

Definition (2.1). G is σ-compact if it has a countable covering by compact subsets.

Definition (2.2). $A[B] = \bigcap_{b \in B} A b^{-1} = \{g \mid g B \in A\}$.

Definition (2.3). $\nu(\cdot)$ is a right-invariant Haar measure on the Borel field generated by the open sets of G.

Sets will be understood to be measurable when so required. Some of the conditions to be considered have been stated in a slightly more restrictive form than necessary in order that they do not involve unmeasurable sets. These restrictions could be lifted, but we feel that the resulting prolongation of the paper would not be justified.

The conditions and their variants are:

(A) ([17], p. 45): For each compact C, there exists sequence $\{C_n\}$, C_n compact, such that $\nu(C_n)/\nu(C_n C) \to 1$.

(A_1): There exists $\{C_n\}$, C_n compact, such that $\nu(C_n)/\nu(C_n C) \to 1$ for all compact C.

(A_2): There exists $\{G_n\}$, G_n closed, $\nu(G_n) < \infty$, such that $\nu(G_n)/\nu(G_n C) \to 1$ for all compact C.

(A_3): For each compact C, there exists $\{G_n\}$, G_n closed, $\nu(G_n) < \infty$, such that $\nu(G_n)/\nu(G_n C) \to 1$.

(A_4): There exists $\{G_n\}$, $\nu(G_n) < \infty$, such that, for each compact C, there is $\{\Gamma_n\}$, $\nu(\Gamma_n) < \infty$, $\Gamma_n \supset G_n C$, such that $\nu(G_n)/\nu(\Gamma_n) \to 1$.

(C) ([2]): There exists $\{N_t \mid t > 0\}$, N_t a compact-open symmetric neighbourhood of the identity e, such that $N_t N_s \subset N_{t+s}$ and $\nu(N_{2t}) < \alpha \nu(N_t)$ for some α independent of t. (Calderon uses left-invariant Haar measure. However, at the present level of abstraction, this distinction is immaterial.)

(FW) ([14]): For each finite F, there exists $\{C_n\}$, C_n compact, such that $\nu(C_n[\{e\} \cup F])/\nu(C_n) \to 1$.

(FW_1): For each finite F, there exists $\{G_n\}$, $\nu(G_n) < \infty$, such that

$$\nu(G_n[\{e\} \cup F])/\nu(G_n) \to 1.$$

(GR) ([10], p. 104): The constant 1 can be approximated uniformly on every compact subset by positive definite functions vanishing outside compact sets.

(H) ([27]): For each compact C, there exists $\{G_n\}$, G_n closed, $\nu(G_n) < \infty$, such that $\nu(G_n[C])/\nu(G_n) \to 1$.

(H_1): There exists $\{G_n\}$, G_n closed, $\nu(G_n) < \infty$, such that $\nu(G_n[C])/\nu(G_n) \to 1$ for all compact C.

(H_2): For each pair $g_1, g_2 \in G$, there exists $\{G_n\}$, $\nu(G_n) < \infty$, such that

$$\nu(G_n[\{e, g_1, g_2\}])/\nu(G_n) \to 1.$$

(H_3): For each compact C, there exists $\{G_n\}$, $\nu(G_n) < \infty$, such that

$$\nu(G_n[\{e, c\}])/\nu(G_n) \to 1$$

uniformly for $c \in C$.

(HS) ([18], p. 336): There is $\{P_n\}$, P_n a probability measure on the Borel sets in G, such that $|P_n(Bg) - P_n(B)| \to 0$ for all $g \in G$ and all B.

(J) ([14] and [19]): There exists a mean m on $L_\infty(G)$ such that $m(f * \lambda) = m(\lambda)$ for any $\lambda \in L_\infty(G)$ and any probability density function (with respect to ν) f.

(LHS) ([18], p. 337): There exists $\{G_n\}$, $\nu(G_n) < \infty$, $G_n \nearrow G$, such that $|P_n(Bg) - P_n(B)| \to 0$ for all $g \in G$ and B, where $P_n(\cdot) = \nu(\cdot \cap G_n)/\nu(G_n)$.

(LHS_1): There exists $\{G_n\}$, $\nu(G_n) < \infty$, such that $|P_n(Bg) - P_n(B)| \to 0$ for all $g \in G$ and B, where $P_n(\cdot) = \nu(\cdot \cap G_n)/\nu(G_n)$.

(M) ([25]): There exists a left-invariant mean on $L_\infty(G)$.

(P) ([4, 25]): For each compact C, there exists $\{f_n\}$, f_n a probability density function (with respect to ν), such that $\int |f_n(gc) - f_n(g)| \, d\nu(g) \to 0$ uniformly for $c \in C$.

(P_1): There exists $\{f_n\}$, f_n as in (P), such that $\int |f_n(gg') - f_n(g)| \, d\nu(g) \to 0$ for all $g' \in G$ and the convergence is uniform on any compact C.

(P_2): There exists $\{f_n\}$, f_n as in (P), such that $\int |f_n(gg') - f_n(g)| \, d\nu(g) \to 0$ for all $g' \in G$.

(P_3): For each pair $g_1, g_2 \in G$, there exists $\{f_n\}$, f_n as in (P), such that

$$\int |f_n(gg_i) - f_n(g)| \, d\nu(g) \to 0, \ i = 1, 2.$$

(Π) ([22]): There exists sequence $\{S_i\}$, $\lim S_i = G$, such that, for each i, there are $\{G_n(i), \Gamma_n(i)\}$, $\nu(G_n(i)) < \infty$, $\nu(\Gamma_n(i)) < \infty$, such that $\Gamma_n(i) \subset G_n(i)[S_i]$ and $\nu(\Gamma_n(i))/\nu(G_n(i)) \to 1$ as $n \to \infty$.

(Π_1): There exists $\{G_n\}$, $\nu(G_n) < \infty$, such that $\nu(G_n[\{e, g\}]/\nu(G_n) \to 1$ for all $g \in G$.

(Π_2): There exists $\{O_i\}$, O_i open $O_i \nearrow G$ such that, for each i, there exists $\{G_n(i)\}$, $G_n(i)$ closed, $\nu(G_n(i)) < \infty$, such that $\nu(G_n(i)[O_i])/\nu(G_n(i)) \to 1$ as $n \to \infty$.

(R) ([8], [28]): The constant 1 can be approximated uniformly on every compact subset by convolutions of the form $x * x^\sim(g)$ where $x(g)$ vanishes outside a compact set and $x^\sim(g) = \overline{x(g^{-1})}$.

The following implication diagram shows how the above conditions are known to be connected. A single arrow means that the implication is proved for σ-compactness of G, except for the case (C) $\to$ (A_4) when connectedness of G is assumed.

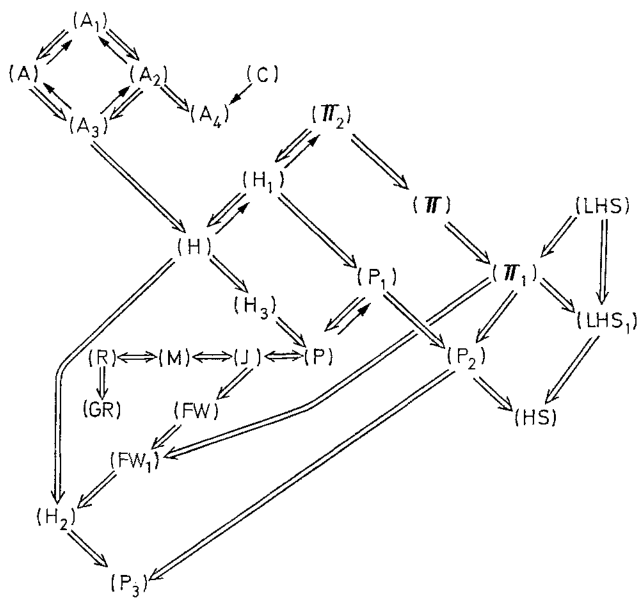

Fig. 1. Implication diagram

Proofs of Implication. (A) $\Rightarrow$ (A_3), (A_1) $\Rightarrow$ (A), (A_1) $\Rightarrow$ (A_2), (A_2) $\Rightarrow$ (A_3), (A_2) $\Rightarrow$ (A_4), (FW) $\Rightarrow$ (FW_1), (FW_1) $\Rightarrow$ (H_2), (H_1) $\Rightarrow$ (H), (H) $\Rightarrow$ (H_2), (LHS) $\Rightarrow$ (LHS_1), (LHS_1) $\Rightarrow$ (HS), (P_1) $\Rightarrow$ (P), (P_1) $\Rightarrow$ (P_2), (P_2) $\Rightarrow$ (P_3), (Π_2) $\Rightarrow$ (Π), (R) $\Rightarrow$ (GR) are immediate.

(A_3) $\Rightarrow$ (H). Given $\{G_n\}$ for C in (A_3), define $G_n^* = G_n C$. If $g \in G_n$, then $g\,C \subset G_n^*$ whence $G_n \subset G_n^*[C]$. Then

$$\frac{\nu(G_n)}{\nu(G_n C)} \leqq \frac{\nu(G_n^*[C])}{\nu(G_n^*)} \leqq 1\,.$$

Take $\{G_n^*\}$ to be the $\{G_n\}$ for (H).

(H) $\Rightarrow$ (H$_3$). Trivial because $\nu(G_n[C]) \leqq \nu(G_n[\{e, c\}]) \leqq \nu(G_n)$.

(H$_3$) $\Rightarrow$ (P). Given $\{G_n\}$ for C in (H$_3$), define $f_n = \chi_{G_n}/\nu(G_n)$.

(H$_2$) $\Rightarrow$ (P$_3$). As for (H$_3$) $\Rightarrow$ (P).

(Π_2) $\Rightarrow$ (H$_1$). Choose $\{\varepsilon_i\}$, $\varepsilon_i \to 0$. For each i, choose n_i such that

$$\nu(G_{n_i}(i)[O_i])/\nu(G_{n_i}(i)) \geqq 1 - \varepsilon_i .$$

For (H$_1$), defining $G_i = G_{n_i}(i)$ suffices since each C is contained in O_j for j large enough.

(H$_1$) $\Rightarrow$ (P$_1$). Given C from (P$_1$), take C in (H$_1$) to be $\{e\} \cup C$. Define $f_n = \chi_{G_n}/\nu(G_n)$.

(Π) $\Rightarrow$ (Π_1). Choose $\{\varepsilon_i\}$, $\varepsilon_i \to 0$. For each i, choose n_i such that

$$\nu(\Gamma_{n_i}(i))/\nu(G_{n_i}(i)) \geqq 1 - \varepsilon_i .$$

For (Π_1), defining $G_i = G_{n_i}(i)$ suffices since, for any $g \in G$, $\Gamma_{n_i}(i) \subset G_i[\{e, g\}]$ for i large enough.

(LHS) $\Rightarrow$ (Π_1). If G is compact, (Π_1) holds trivially. For G non-compact, suppose $\{G_n\}$ given by (LHS). Since $G_n \nearrow G$, $\nu(G_n) \to \infty$. Without loss of generality, suppose $\nu(G_n)/\nu(G_{n+1}) \leqq \frac{1}{2}$, $n = 1, 2, \ldots$. Choose g for (Π_1). Let

$$A_n = G_n - G_n g^{-1}, \qquad W_n = \bigcup_{i=1}^{n} A_i, \qquad B = G - W_\infty .$$

Then $(Bg) \cap G_n = G_n - W_\infty g$. But $A_i g \subset G - G_i$. So

$$W_\infty g \cap G_n = \left(\bigcup_{i=1}^{n-1} A_i g\right) \cap G_n = W_{n-1} g \cap G_n ,$$

whence $(Bg) \cap G_n = G_n - W_{n-1} g$. So $\nu[(Bg) \cap G_n] \geqq \nu(G_n) - \nu(W_{n-1})$. But

$W_{n-1} \subset G_{n-1} \cap W_\infty = G_{n-1} - B \cap G_{n-1}$ so that

$\nu[(Bg) \cap G_n] \geqq \nu(G_n) - [\nu(G_{n-1}) - \nu(B \cap G_{n-1})]$ or

$$\nu_n(Bg) \geqq 1 - \frac{\nu(G_{n-1})}{\nu(G_n)} [1 - \nu_{n-1}(B)] \geqq \tfrac{1}{2} + \tfrac{1}{2} \nu_{n-1}(B) . \tag{2.1}$$

But, by (LHS), $|\nu_n(Bg) - \nu_n(B)| \to 0$. That is, given $\varepsilon > 0$, there exists $n(\varepsilon)$ such that, for $n > n(\varepsilon)$, $\nu_n(Bg) < \nu_n(B) + \varepsilon$. Whence, by (2.1), $\nu_n(B) + \varepsilon > \frac{1}{2} + \frac{1}{2} \nu_{n-1}(B)$ for $n > n(\varepsilon)$, repeated use of which inequality yields

$$\nu_{n(\varepsilon)+m}(B) > (1 - 2\varepsilon)(2^{-1} + \cdots + 2^{-m}) + 2^{-m} \nu_{n(\varepsilon)}(B) .$$

So $\liminf \nu_n(B) \geqq 1 - 2\varepsilon$. But ε is arbitrary, therefore $\nu_n(B) \to 1$. But $B \cap G_n = G_n - W_\infty \subset G_n - A_n = G_n \cap (G_n g^{-1})$. Whence $\nu(G_n[\{e, g\}])/\nu(G_n) \to 1$.

(Π_1) $\Rightarrow$ (P$_2$). Define $f_n = \chi_{G_n}/\nu(G_n)$.

(Π_1) $\Rightarrow$ (LHS$_1$). Using $f_n = \chi_{G_n}/\nu(G_n)$, it is readily seen that the $\{G_n\}$ in (Π_1) suffice for (LHS$_1$).

(Π_1) $\Rightarrow$ (FW$_1$). The $\{G_n\}$ in (Π_1) suffice for all F of (FW$_1$) because

$$G_n[\{e\} \cup F] = G_n - \bigcup_{f \in F} (G_n - G_n f^{-1})$$

and $\nu(G_n - G_n f^{-1})/\nu(G_n) \to 0$ by (Π_1).

$(P_2) \Rightarrow (HS)$. Define $\nu_n(B) = \int_B f_n(g)\, d\nu(g)$.

$(A_3) \to (A)$. G σ-compact implies $G = \lim K_i$, K_i compact. Given $\{G_n\}$ for C in (A_3), define $C_{in} = K_i \cap G_n$. Then C_{in} is compact and $\nu(C_{in}) \to \nu(G_n)$. Choose $\{\varepsilon_n\}$, $\varepsilon_n \to 0$ and take $i(n)$ such that $|\nu(C_{i(n)n}) - \nu(G_n)| < \varepsilon_n \nu(G_n)$. Then $C_n = C_{i(n)n}$ will suffice for (A).

$(A_2) \to (A_1)$. As for $(A_3) \to (A)$.

$(A_3) \to (A_2)$. G σ-compact implies $G = \lim O_i$, O_i open, with $\bar{O}_i = C_i$ compact and $C_i \subset O_{i+1}$. By (A_3), there exists $\{G_n(i)\}$ such that $\nu(G_n(i))/\nu(G_n(i)C_i) \to 1$ as $n \to \infty$. Choose $\{\varepsilon_i\}$, $\varepsilon_i \to 0$ and take $n(i)$ such that

$$\nu(G_{n(i)}(i))/\nu(G_{n(i)}(i)\, C_i) > 1 - \varepsilon_i .$$

Then $G_n = G_{n(i)}(i)$ will suffice for (A_2) because any compact C is included in O_i, and therefore in C_i, for i large enough and $G_i C \subset G_i C_i$, whence $\nu(G_i)/\nu(G_i C) > 1 - \varepsilon_i$.

$(A) \to (A_1)$. As for $(A_3) \to (A_2)$.

$(H) \to (H_1)$. As for $(A_3) \to (A_2)$.

$(H_1) \to (\Pi_2)$. The $\{O_i\}$ in $(A_3) \to (A_2)$ suffice, because, by (H_1), there exists $\{G_n\}$, G_n closed, such that $\nu(G_n[C_i])/\nu(G_n) \to 1$ as $n \to \infty$.

$(P) \to (P_1)$. Define O_i and C_i as in $(A_3) \to (A_2)$.

By (P), $\exists \{f_{ni}\}$ such that

$$\int |f_{ni}(gc) - f_{ni}(g)|\, d\nu(g) \to 0 \quad \text{uniformly}$$

for $c \in C_i$. Choose $\{\varepsilon_i\}$, $\varepsilon_i \to 0$ and take $n(i)$ such that

$$\int |f_{n(i)i}(gc) - f_{n(i)i}(g)|\, d\nu(g) < \varepsilon_i \quad \text{for} \quad c \in C_i .$$

Then $f_i = f_{n(i)i}$ suffices for (P_1).

$(C) \to (A_4)$. By [2], (C) implies there exists $\{t_n\}$, $t_n \to \infty$, such that, for every $s \geqq 0$, $\nu(N_{t_n})/\nu(N_{s+t_n}) \to 1$ as $n \to \infty$. Since N_1 is a neighbourhood of e there exists an open neighbourhood O of e, $O \subset N_1$. By the assumption that G is connected and p. 62 of [12], $G = \lim O^i$. So, given compact C, there exists p such that $C \subset O^p \subset N_1{}^p \subset N_p$, the latter step by (C). Defining $G_n = N_{t_n}$ and $\Gamma_n = N_{p+t_n}$ suffices for (A_4) since $G_n C \subset N_{t_n} N_p \subset \Gamma_n$ and $\nu(N_{t_n})/\nu(N_{p+t_n}) \to 1$ as $n \to \infty$.

For $(J) \Rightarrow (FW)$, see p. 99 of [14]. For $(M) \Leftrightarrow (P)$, see [25].

For $(P) \Leftrightarrow (R)$, see [24]. For $(J) \Leftrightarrow (P)$, see [14].

For $(M) \Rightarrow (J)$, see [19].

Emerson and Greenleaf [29] have recently shown that $(H_3) \Rightarrow (A)$ for any locally compact topological group.

§ 3. Counterexamples

We firstly state for (P_3), the counterpart of the theorem proved by Reiter [23] for (P).

Theorem (3.1). *If G has* (P_3) *then every discrete subgroup of G has* (P_3) *also.*
The proof follows Reiter's proof in an obvious manner.

The free discrete group with two generators, F, was used as a counterexample to the vacuity of their conditions by VON NEUMANN [20], DAY [3] and DIEUDONNÉ [5].

Likewise we have:

Theorem (3.2). *F does not obey* (P_3).

Proof. We first show that F has (P_3) implies F has (P). Suppose g_1, g_2 are the free generators of F. Then there exist probability density functions $\{f_n\}$ such that $\sum_{g\in F} |f_n(gg_i) - f_n(g)| \to 0$ as $n \to \infty$, $i = 1, 2$.

An arbitrary compact set C in F is a finite set. Suppose $g' \in C$ and that, in reduced form, $g' = x_1^{\varepsilon_1} x_2^{\varepsilon_2} \ldots x_k^{\varepsilon_k}$, $x_i = g_1$ or g_2, $\varepsilon_i = \pm 1$. Then

$$\begin{aligned}
&\sum_{g\in F} |f_n(gg') - f_n(g)| \\
&= \sum_{g\in F} \left| \sum_{s=1}^{k} [f_n(g x_1^{\varepsilon_1} \ldots x_s^{\varepsilon_s}) - f_n(g x_1^{\varepsilon_1} \ldots x_{s-1}^{\varepsilon_{s-1}})] \right| \\
&\leqq \sum_{s=1}^{k} \sum_{g\in F} |f_n(g x_s^{|\varepsilon_s|}) - f_n(g)| \\
&\to 0 \text{ as } n \to \infty .
\end{aligned}$$

Whence $\{f_n\}$ suffices for (P). So F has (P_3) implies F has (P). But DIEUDONNÉ [5] showed that F does not obey (P).

From Theorems (3.1) and (3.2) it follows that any G containing F as a subgroup does not obey (P_3). Such groups include $SL_n(R)$ and $GL_n(R)$, $n \geqq 2$. This result clearly relates to (a) KUDO's suggestion [17] that $GL_n(R)$, $n \geqq 2$, did not have (A) (b) STEIN's statistically based demonstration (p. 338 of [18]) that $GL_n(R)$, $n \geqq 2$, could not have HS (c) the statistically based demonstrations of KIEFER [16] and STEIN [21], § 5, of the non-vacuity of (Π_1) and (A).

The proofs of most hitherto published counterexamples to vacuity have been indirect, that is, have rested on some extensive theory, either group theoretical or statistical. It is therefore of some interest to exhibit the following simple, direct proof that F does not possess (H_2). Suppose that, with g_1, g_2 the free generators of F, there existed $\{G_n\}$ with $\nu(G_n) < \infty$ such that

$$\nu(G_n \cap G_n g_i^{-1})/\nu(G_n) \to 1\,, \quad i = 1, 2\,. \tag{3.1}$$

Since F is discrete, ν is counting measure so that G_n is finite. Suppose G_n has $n(1, 1)$ elements ending in g_1, $n(1, -1)$ ending in g_1^{-1}, $n(2, 1)$ ending in g_2 and $n(2, -1)$ ending in g_2^{-1} with $N = n(1, 1) + n(1, -1) + n(2, 1) + n(2, -1)$. Then $G_n g_1^{-1}$ has $n(1, -1) + n(2, 1) + n(2, -1)$ elements ending in g_1^{-1}, while $G_n g_2^{-1}$ has $n(1, 1) + n(1, -1) + n(2, -1)$ elements ending in g_2^{-1}. So, by (3.1),

$$\begin{aligned}
&[n(1, -1) - \{n(1, -1) + n(2, 1) + n(2, -1)\}]/N \to 0 \quad \text{and} \\
&[n(2, -1) - \{n(1, 1) + n(1, -1) + n(2, -1)\}]/N \to 0
\end{aligned}$$

implying the absurdity $N/N \to 0$.

The conjecture in [27] that the group $UT(2)$ of non-singular upper triangular 2×2 matrices does not have (H) is refutable by the demonstration (too extended to reproduce here) that the sequence $\{G_n\}$ with

$$G_n = \left\{ \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \,\middle|\, n^{-1} < |a| < n,\ |b| < \exp n^2,\ |c| > 1,\ |c| > |b| \right\}$$

will suffice for (H_1) and a fortiori for (H).

§ 4. Statistical Implications and Connexions

All of the conditions in § 2 are satisfied when G is compact. So, in what follows, it will be understood that G is non-compact.

An account of the statistical relevance of the conditions (HS) and (LHS) is given in Chapter 8 of [18]. The probability distributions $\{P_n\}$ in these conditions constitute an asymptotically right invariant sequence that induces a similar sequence of what may be considered prior distributions on Θ. The existence of such a sequence has the consequence that, for certain testing problems defined for the invariant experiment, there is a minimax almost invariant test.

Peisakoff [22] showed that (Π_1) and (Π) are involved in proofs of the ε-minimaxity of the class of invariant decision functions, as well as in his corollary (2.7), which gives conditions for the minimaxity of the (invariant) quasi Bayes decision function generated by right Haar measure as quasi prior distribution. The latter property is also proved by Kudo [17] under the assumption of (A). Kiefer adopts (Π_1) in [16] for similar purposes. Stone [27] invoked condition (H) as a sufficient condition for a certain convergence justification of the use of right Haar measure as quasi prior distribution. Grenander [10], pp. 121 and 126, involves (GR) in connexion with applications of certain probabilistic limit theorems.

We will now state three further statistical connexions:

(a) *First statistical interpretation of* (P_3) *and its negation.* Consider a sequence of experiments $\{\mathscr{E}_n\}$, all referring to the same data space G, a locally compact topological group. For all $\mathscr{E}_n$, there is the same statistical problem, inference about a parameter θ with three possible values represented by hypotheses:

$$\begin{aligned} H_0&\colon\ \theta = e, && \text{the identity element of } G,\\ H_1&\colon\ \theta = g_1, && g_1 \in G,\\ H_2&\colon\ \theta = g_2, && g_2 \in G. \end{aligned}$$

For each $\mathscr{E}_n$, the problem is of location parameter type with a probability density function. That is, we suppose that, under H_0, H_1, H_2, g has probability density functions $f_n(g)$, $f_n(gg_1)$, $f_n(gg_2)$, respectively, with respect to right Haar measure. When G has (P_3), we know that there exist $\{f_n\}$ such that

$$\int |f_n(gg_i) - f_n(g)|\, d\nu(g) \to 0 \tag{4.1}$$

$i = 1, 2$. The statistical implication of (4.1) is that the experiments $\{\mathscr{E}_n\}$ are *asymptotically uninformative* as $n \to \infty$. For instance, suppose we were to test H_i against H_0 with critical (rejecting H_0) region C_{in} and errors α_{in}, β_{in} of the first and second kinds respectively. Then

$$\begin{aligned}\alpha_{in} + \beta_{in} &= \int_{C_{in}} f_n(g)\,d\nu(g) + \Big(1 - \int_{C_{in}} f_n(g g_i)\,d\nu(g)\Big)\\ &= 1 - \int_{C_{in}} [f_n(g g_i) - f_n(g)]\,d\nu(g)\\ &\to 1\end{aligned}$$

by (4.1). But $\alpha + \beta = 1$ is the characteristic of a completely uninformative test.

On the other hand, if G does not have (P_3), there do not exist $\{f_n\}$ such that $\{\mathscr{E}_n\}$ are asymptotically uninformative. For, if (4.1) is not satisfied, there exists $\lambda > 0$ such that for $i = i^*$ and for the subsequence $\{\mathscr{E}_{n^*}\}$

$$\int |f_{n^*}(g g_{i^*}) - f_{n^*}(g)|\,d\nu(g) > \lambda\,.$$

For $\mathscr{E}_{n^*}$, the test of H_{i^*} against H_0 using $C_{n^* i^*} = \{g \mid f_{n^*}(g g_{i^*}) > f_{n^*}(g)\}$ has

$$\alpha_{i^* n^*} + \beta_{i^* n^*} = 1 - \int_{C_{n^* i^*}} [f_{n^*}(g g_{i^*}) - f_{n^*}(g)]\,d\nu(g)\,.$$

But

$$\lambda < 2 \int_{C_{n^* i^*}} [f_{n^*}(g g_{i^*}) - f_{n^*}(g)]\,d\nu(g)\,.$$

Therefore $\alpha_{i^* n^*} + \beta_{i^* n^*} < 1 - \frac{1}{2}\lambda$, showing that the sequence $\{\mathscr{E}_{n^*}\}$ is not asymptotically uninformative.

(b) *Second statistical interpretation of* (P_3) *and its negation.* Suppose (P_3) holds. If $x(g)$ is a bounded random variable then for g_1, g_2 and associated $\{f_n\}$

$$\begin{aligned}&\Big|\int f_n(g)\,x(g g_i^{-1})\,d\nu(g) - \int f_n(g)\,x(g)\,d\nu(g)\Big|\\ &= \Big|\int [f_n(g g_i) - f_n(g)]\,x(g)\,d\nu(g)\Big|\\ &\leq \int |f_n(g g_i) - f_n(g)|\,d\nu(g)\,.\,\sup_G x(g)\\ &\to 0\end{aligned} \tag{4.2}$$

as $n \to \infty$. If we now interpret $\{f_n\}$ as a sequence of prior probability density functions on G then (4.2) states that the difference between the prior expectations of $x(g)$ and its translate through g_i tends to zero for $i = 1, 2$.

Now the latter must be regarded as a natural requirement for $\{f_n\}$ to be regardable as an asymptotically uninformative sequence of priors. That is we would not expect the expectation of any (bounded) random variable to be much affected by the translations g_i, $i = 1, 2$. On the other hand it is clear that if (P_3) does not hold then there exist g_1, g_2 such there is not a sequence of such asymptotically uninformative priors with respect to all bounded random variables.

(c) *A wide sense Bayes link.* Adopting the conventional formulation of (i) an observation x in some general space, whose distribution with density $f(x \mid \theta)$ depends on a general parameter $\theta \in \Theta$ (ii) a general decision space $D = \{d\}$ and a real-valued loss function $W(d, \theta)$ (iii) the representation of non-randomized decision function $d(x)$ by δ.

Definition (4.1). With respect to some metric $M(d_1, d_2)$ we have "convergence in probability to decision function δ" if there is a sequence of proper prior measures

$\{\pi_n\}$ on Θ with corresponding strict Bayes decision functions $\{\delta_n\}$ such that

$$\operatorname*{plim}_{n\to\infty} M[d(x), d_n(x)] = 0 \tag{4.3}$$

where, in the evaluation of plim, x is given, at each n, its marginal distribution

$$\int f(x\,|\,\theta)\, d\pi_n .$$

Theorem (4.1). *If $M(d_1, d_2) = \sup_\theta (W(d_1, \theta) - W(d_2, \theta)|$ is bounded then convergence in probability to δ with respect to M implies δ is wide sense Bayes.*

Proof. Let $R(\pi, \delta)$ denote the Bayes risk of δ with respect to prior π. Then

$$R(\pi_n, \delta) - R(\pi_n, \delta_n) \leqq E\, M[d(x), d_n(x)] .$$

So, if M is bounded,

$$\operatorname*{plim}_{n\to\infty} M[d(x), d_n(x)] = 0 \Rightarrow E\, M[d(x), d_n(x)] \to 0$$

or δ is wide-sense Bayes with respect to $\{\pi_n\}$.

The relevance of Definition (4.1) and Theorem (4.1) is that Kudo [17] showed that, under certain conditions which included (A), the (invariant) right Haar decision function, for the invariant experiment in which G is isomorphic to parameter space Θ, is wide-sense Bayes. However we may think of convergence in probability to decision functions as a generalization of the definition of convergence in probability of posterior distributions adopted by Stone [27] for which condition (H) was invoked. We thus obtain some insight into the need for conditions (A) and (H) in apparently different areas of statistics.

Acknowledgement. The paper is a radical improvement of an earlier version by the first author thanks mainly to some helpful references from a referee.

References

1. Brillinger, D. R.: Necessary and sufficient conditions for a statistical model to be invariant under a Lie group. Ann. math. Statistics **34**, 492—500 (1963).
2. Calderon, A. P.: A general ergodic theorem. Ann. of. Math., II. Ser. **58**, 182—191 (1953).
3. Day, M. M.: Amenable semigroups. Illinois J. Math. **1**, 509—544 (1957).
4. Dieudonné, J.: Sur le produit de composition (II). J. Math. pur. appl., IX. Ser. **39**, 275—292 (1960).
5. — Sur un propriété des groupes libres. J. reine angew. Math. **204**, 30—34 (1960).
6. Fraser, D. A. S.: Statistical models and invariance. Ann. math. Statistics **38**, 1061—1067 (1967).
7. Følner, E.: Note on groups with and without full Banach mean value. Math. Scandinav. **5**, 5—11 (1957).
8. Godement, R.: Les fonctions de type positif et la théorie des groupes. Trans. Amer. math. Soc. **63**, 1—84 (1948).
9. Greenleaf, F. P.: Invariant means on locally compact Groups and their Applications. To appear in the Van Nostrand Mathematical Studies Series. 1967.
10. Grenander, U.: Probabilities on algebraic structures. New York and London: Wiley 1963.
11. Hannan, E. J.: Group representations and applied probability. J. appl. Probab. **2**, 1—68 (1965).
12. Hewitt, E. A., and K. A. Ross: Abstract harmonic analysis. Berlin-Göttingen-Heidelberg: Springer 1963.
13. Hora, R. B., and R. J. Buehler: Fiducial theory and invariant estimation. Ann. math. Statistics **37**, 643—656 (1966).

14. Hulanicki, A.: Means and Følner condition on locally compact groups. Studia math. **27**, 87–104 (1966).
15. Kesten, H.: Full Banach mean values on countable groups. Math. Scandinav. **7**, 146–156 (1959).
16. Kiefer, J.: Invariance, minimax sequential estimation, and continuous time processes. Ann. math. Statistics **28**, 573–601 (1957).
17. Kudo, H.: On minimax invariant estimates of the transformation parameter. Natur. Sci. Rep. Ochanomizu Univ. **6**, 31–73 (1955).
18. Lehmann, E. L.: Testing statistical hypotheses. New York and London: Wiley 1959.
19. Namioka, I.: On a recent theorem by H. Reiter. Proc. Amer. math. Soc. **17**, 1101–1102. (1966).
20. Neumann, J. von: Zur allgemeinen Theorie des Maßes. Fundamenta Math. **13**, 73–116 (1929).
21. Neyman, J.: Proc. Fourth Berkeley Sympos. math. Statist. Probability **1**, 361–379 (1961).
22. Peisakoff, M.: Transformation parameters. Thesis, Princeton University. University Microfilms, Ann Arbor, Michigan, 1950.
23. Reiter, H.: Sur les groupes de Lie semi-simples connexes. C. r. Acad. Sci. Paris **255**, 2883–2884 (1962).
24. — Sur la propriété (P_1) et les fonctions de type positif. C. r. Acad. Sci., Paris **258**, 5134–5135 (1964).
25. — On some properties of locally compact groups. Indagationes math. **27**, 697–701 (1965)
26. Stegeman, J. D.: On a property concerning locally compact groups. Indagationes math. **27**, 701–703 (1965).
27. Stone, M.: Right Haar measure for convergence in probability to quasi posterior distributions. Ann. math. Statistics **36**, 440–453 (1965).
28. Takenouchi, O.: Sur une classe de fonctions continues de type positif sur un groupe localement compact. Math. J. Okayama Univ. **4**, 143–173 (1965).
29. Emerson, W. R., and F. P. Greenleaf: Covering properties and Følner conditions for locally compact groups. Math. Z. **102**, 370–384 (1967).

Dr. M. Stone
Dept. of Statistics
University College London
Gower Street
London, Great Britain

Dr. R. von Randow
Dept. of Mathematics
University of Durham
South Road
Durham, Great Britain