Z. Wahrscheinlichkeitstheorie verw. Gebiete
57, 103–128 (1981)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Amenability: A Survey for Statistical Applications of Hunt-Stein and Related Conditions on Groups

James V. Bondar and Paul Milnes*

Dept. of Mathematics, University of Western Ontario, London, Ontario, Canada N6A5B7

Summary. A number of conditions on groups have appeared in the literature of invariant statistical models in connection with minimaxity, approximation of invariant Bayes priors by proper priors, the relationship between Bayesian and classical inference, ergodic theorems, and other matters. In the last decade, rapid development has occurred in the field and many of these conditions are now known to be equivalent. We survey the subject, make the equivalences explicit, and list some groups of statistical interest which do, and also some which do not, have these properties. In particular, it is shown that the existence of the asymptotically invariant sequence of probabilities in the hypothesis of the Hunt-Stein theorem is equivalent to amenability, a condition that has been much studied by functional analysts.

0. Introduction

In this survey we consider a class of conditions on groups which have arisen in diverse and seemingly unrelated investigations on invariant statistical models, for example, investigations concerning: minimaxity and theorems of the Hunt-Stein type (Peisakoff (1950), Kiefer (1957), Wesler (1959), etc.), approximately least favorable or uninformative prior distributions (Zehnwirth (1975)), ergodic theorems on groups of transformations (see Sect. 2 for references), the spectra of transition matrices for random walks on groups (Kesten (1959)), the relation between Bayes and Neyman-Pearson confidence intervals (Bondar (1977), Heath and Sudderth (1978)), the relationship between proper and improper Bayes procedures (Stein (1965), Stone (1970)), and others.

In all of these papers, the problem can be related to finding some sort of invariant average (probability) over the group; particularly well-behaved averages exist in compact groups, namely, Haar measure. The groups possessing such invariant averages are called amenable groups; thus amenability may be

* Research supported in part by NSERC grant A7857

0044-3719/81/0057/0103/$05.20

viewed as a generalization of compactness. An appealing reason for using amenability in statistics is the remarkable fact that, if the group of a statistical problem has the algebraico-topological property called amenability, the truth of many of the results which interest us depends little on the exact nature of the particular action of the group on the sample space, or on the nature of the particular probability distributions in the problem.

A survey of the interrelations then known to exist among some of these group-theoretic conditions was made by Stone and von Randow (1968). They also conjectured that many of the conditions were equivalent. During the years following the writing of their paper, the subject received much attention from pure mathematicians, and many equivalences among these conditions were proved in numerous papers by various authors, scattered through sundry journals; a good number of these was collected in Greenleaf (1969). Other equivalences are known but seem not to be in the literature.

Our main theorem (end of Sect. 1) will make the equivalences explicit. In Sect. 2, we collect some (known) results partially characterizing the groups which possess the properties being discussed. Some proofs are given in Sect. 3, and in the appendix is assembled what seems, to the authors at least, a more direct and up-to-date exposition than has yet appeared in the literature, of the chain of implications from the seemingly weaker condition that the bounded uniformly continuous functions on G have a right invariant mean, to the apparently stronger condition, that G has a summing net.

For a discussion of the statistical applications, the reader is referred to the last half of Sect. 2, where among other things, a proof of the Hunt-Stein theorem is given.

1. The Conditions and Their Equivalences

In the following, G will be a locally compact, Hausdorff topological group; ν will be a right Haar measure on G and μ the associated left Haar measure ($\mu(E)=\nu(E^{-1})$, where $E^{-1}=\{g^{-1}|g\in E\}$); the Borel sets will be those generated by the open sets; e denotes the identity of G. See for example, Nachbin (1965), Chap. 2 for definitions. In Hewitt and Ross (1963), Sect. 15.10, 15.27 and 20.32 contain useful remarks about the relations between left and right Haar measures. We will consider only real-valued functions on G; most arguments we use will apply directly to complex-valued functions. Unless specific mention is made to the contrary, probability density functions are assumed to be with respect to ν (w.r.t. ν), which means they are in $L_1(G,\nu)$, which in general is different from $L_1(G,\mu)$. Null sets will be those whose ν-measure is zero (this is equivalent to having μ-measure of zero, hence $L_\infty(G,\nu)=L_\infty(G,\mu)(=L_\infty(G)$, say)).

If $\mathscr{V}$ is a vector space of essentially bounded, measurable (real-valued) functions on G, which contains the constant function $\mathbf{1}$, a *mean* on $\mathscr{V}$ is a linear functional m on $\mathscr{V}$ such that

(i) $m(\mathbf{1})=1$ and $m(f)\geqq 0$ if $f\geqq 0$ a.e. (ν).

It follows that $|m(f)| \leqq \operatorname*{ess\,sup}_{g \in G} |f(g)|$, i.e., m is continuous. Also (i) is equivalent to

(ii) $\operatorname*{ess\,inf}_{g \in G}(g) \leqq m(f) \leqq \operatorname*{ess\,sup}_{g \in G} f(g)$ for all $f \in \mathscr{V}$.

The set $\mathscr{M} = \mathscr{M}(\mathscr{V})$ of all means on $\mathscr{V}$ is compact in the weak* topology (i.e., $\sigma(\mathscr{V}^*, \mathscr{V})$-compact); this is the topology on $\mathscr{M}$ which has as subbase the family of all sets of the form $\{m \in \mathscr{M} \mid |m(f) - m_0(f)| < a\}$, where $f \in \mathscr{V}$, $a > 0$ and $m_0 \in \mathscr{M}$ (see Example 1 below for an illuminating special case). The probability densities on G are weak*-dense in $\mathscr{M}$; and in case $\mathscr{V} \subset C(G)$, the space of bounded continuous functions on G, then convex combinations of the *evaluation means* $\{p_g \mid g \in G\}$, where $p_g(f) = f(g)$ for $f \in C(G)$ and $g \in G$, are weak*-dense in $\mathscr{M}$. The *right translate* f_g (resp. *left translate* ${}_gf$) of a function f on G by $g \in G$ is defined by $f_g(g') = f(g'g)$ (resp. ${}_gf(g') = f(gg')$) for all $g' \in G$. A vector space $\mathscr{V}$ of functions on G is called *right* (resp. *left*) *translation invariant* if $f_g \in \mathscr{V}$ (resp. ${}_gf \in \mathscr{V}$) whenever $f \in \mathscr{V}$, $g \in G$; and a mean m on $\mathscr{V}$ is called *right* (resp. *left*) invariant provided $m(f_g) = m(f)$ (resp. $m({}_gf) = m(f)$) for all $f \in \mathscr{V}$, $g \in G$. When we say that a mean, vector space or what-have-you is *invariant*, we mean that it is simultaneously both left and right invariant. If $f \in L_\infty(G) = L_\infty$ and $p \in L_1(G, \nu) = L_1$, then the convolutions $f * p$ and $\check{p} * f$ are defined almost everywhere (w.r.t. ν) by

$$(f * p)(s) = \int_G f(sg^{-1})\, p(g)\, d\nu(g)$$

and

$$(\check{p} * f)(s) = \int_G p(gs^{-1})\, f(g)\, d\nu(g)$$

(where $\check{p}(t) = p(t^{-1})$, $t \in G$), and are in $\mathrm{LUC}(G) = \mathrm{LUC}$, and $\mathrm{RUC}(G) = \mathrm{RUC}$ respectively, with $\|f * p\|_\infty \leqq \|f\|_\infty \|p\|_1$ and $\|\check{p} * f\|_\infty \leqq \|p\|_1 \|f\|_\infty$. RUC (resp. LUC) is the space of bounded right (resp. left) uniformly continuous functions on G. (A function f on G is *right* (resp. *left*) *uniformly continuous* if, given any $\varepsilon > 0$, there is a neighbourhood $V = V(\varepsilon)$ of $e \in G$ such that $|f(s) - f(t)| < \varepsilon$ whenever $st^{-1} \in V$ (resp. $t^{-1}s \in V$).) And finally, if $\mathscr{V}$ is a subspace of L_∞ containing the constant functions and such that $f * p \in \mathscr{V}$ (resp. $\check{p} * f \in \mathscr{V}$) whenever $f \in \mathscr{V}$ and $p \in L_1$, a mean m on $\mathscr{V}$ is called *topologically right* (resp. *left*) *invariant* if $m(f * p) = m(f)$ (resp. $m(\check{p} * f) = m(f)$) whenever $f \in \mathscr{V}$ and $p \in L_1$ satisfies

$$\|p\|_1 = \int_G p(g)\, d\nu(g) = 1. \tag{$*$}$$

(The members of L_1 that satisfy condition $(*)$ are precisely the probability densities; we remind the reader that all probability densities are with respect to ν unless specific mention is made to the contrary.)

Example 1. If R is the usual additive real numbers and $n \in N$, then the equation

$$m_n(f) = (2n)^{-1} \int_{-n}^{n} f(x)\, dx$$

defines a mean m_n on $L_\infty(R)=L_\infty$. By weak* compactness of $\mathscr{M}=\mathscr{M}(L_\infty)$, the sequence $\{m_n\}$ has a weak* accumulation point m in $\mathscr{M}$, i.e., m is a mean on L_∞ such that, given $n_0 \in N$, $\varepsilon>0$ and $f_1, f_2, \ldots, f_k \in L_\infty$, there is an $n \geqq n_0$ with

$$|m(f_i)-m_n(f_i)|<\varepsilon, \quad i=1,2,\ldots,k.$$

It is easy to verify that m is a (topologically right) invariant mean. We note three things about m (all of which also hold for invariant means on general non-compact, σ-compact, locally compact groups):

(i) $m(f)=0$ if $f(x)\to 0$ as $x\to\infty$.

(ii) m is not σ-additive in the sense that, if f_i is defined for $i\in N$ by

$$f_i(x)=\begin{cases}1 & \text{if } -i\leqq x<-i+1 \text{ or } i-1\leqq x<i\\ 0 & \text{otherwise,}\end{cases}$$

then $m\left(\sum_1^\infty f_i\right)=1 \neq 0=\sum_1^\infty m(f_i)$. Thus m determines a probability measure on the Borel subsets of R that is finitely additive and not σ-additive. See Dunford and Schwartz (1958), Theorem IV.5.1, p. 258, in this regard. (Of course, m cannot be invariant and σ-additive, because this would imply it was a multiple of Haar measure, i.e., Lebesgue measure.)

(iii) m is not the only right invariant mean on L_∞ or, what is the same thing, $\{m_n\}$ does not converge to m in the weak* topology. For example, if $\{a_j\}$ is a sequence of positive numbers such that $(a_{j+1}-a_j)/a_j\to\infty$ as $j\to\infty$, and $F\in L_\infty$ is defined by

$$F(x)=\begin{cases}1 & \text{if } -a_{2j+1}\leqq x<-a_{2j} \text{ or } a_{2j}\leqq x<a_{2j+1} \text{ for } j\in N\\ 0 & \text{elsewhere,}\end{cases}$$

then $\liminf_n m_n(F)=0$, $\limsup_n m_n(F)=1$, and $\{m_n(F)\}$ does not converge. By doing a similar construction, one can show that no subsequence of $\{m_n\}$ converges to m either. In fact, it is always the case for a non-compact locally compact group with an invariant mean m on L_∞ that no sequence of probability densities converges to m in the weak* topology; thus, even in the case of R one must use nets that are not sequences: the subnet of $\{m_n\}$ that converges to m is not a subsequence. (We note that the cardinality of the set of invariant means is very large indeed; see Chou (1970, 1976) and Remark (iv) following condition J ahead.)

Any group currently of interest in parametric statistics is *σ-compact*, i.e., the group is a countable union of compact subsets. If G is σ-compact and satisfies any of the conditions below which are stated in terms of nets of sets or measures, then the group satisfies the same condition with the net replaced by a sequence (and conversely, of course, since any sequence is a net). Thus, those who are only interested in σ-compact groups can read "sequence" for "net" in the definitions and the statements of theorems below. The use of nets in the general case is necessary for, unless a group is σ-compact, it cannot satisfy many of the conditions involving sequences (Emerson (1968), Theorem 4). If G

is not σ-compact one must take care to define L_∞ properly (functions are to be identified if they differ only on *locally* null sets); see Greenleaf (1969), p. 22, for technical comments.

Each condition given below occurs in both a left-handed and a right-handed form; the left-handed form is just the right-handed condition with μ replacing ν, left multiplication replacing right multiplication and ${}_gf$ replacing f_g. The mapping $g \to g^{-1}$ on G will map sets and measures satisfying a right-handed condition into ones satisfying the corresponding left-handed condition and vice versa, showing that G satisfies a right-handed condition if and only if it satisfies the analogous left-handed condition. Of course, for commutative groups the left- and right-handed forms are identical in content.

We adopt what we consider to be the major conditions of Stone and von Randow (1968) and also much of their nomenclature. We use nets rather than sequences in order to cover the non-σ-compact case; Stone and von Randow do not use nets explicitly. All conditions will be put in their right-handed forms. The reader is warned that, in the literature of pure mathematics, it is customary to use the left-handed forms; no confusion should result in view of the above remarks on the equivalence of left- and right-handed versions.

We now define our conditions. The rationale for the order in which we give them would be: while also making some attempt to keep similar conditions together, we have started with the conditions of the Appendix, in the order in which they appear there, and have then given the other conditions of statistical interest. The theorem relating these conditions to each other is stated at the end of this section. A discussion of their statistical applications can be found at the end of Sect. 2.

Condition M (amenability) – there exists a right invariant mean on $L_\infty(G)$. This is known (Greenleaf (1969), Theorem 2.2.1) to be equivalent to: M_{cb} – there exists a right invariant mean on the space $C(G)$ of bounded continuous functions on G; and also equivalent to: M_{ucb} – there exists a right invariant mean on the space $\mathrm{LUC} \cap \mathrm{RUC} = \mathrm{UC}$.

By our previous remarks, M, M_{cb} and M_{ucb} are equivalent to their left-handed forms, namely the existence of left invariant means on the function spaces L_∞, $C(G)$ and UC, respectively. In fact, if there is a right invariant mean, then there is a mean which is simultaneously left *and* right invariant (Greenleaf (1969), p. 29, or see Remark (iii) below.

Condition J (topological amenability) – there exists a topologically right invariant mean m on $L_\infty(G)$. This condition is known (Greenleaf (1969), Sect. 2.2) to be equivalent to J_{cb} (resp. J_{ucb}), the condition obtained by replacing $L_\infty(G)$ by $C(G)$ (resp. UC).

Remarks. (i) If p_g is the evaluation mean at g, i.e., the probability measure whose mass is concentrated at g, then $f_{g^{-1}} = f * p_g$ (see Remark (ii) for the definition of $f * p_g$), so we see that J can be considered as a "topological" version of the "discrete" condition M. (The left-handed form of J is $m(p *_\mu f) = m(f)$ for every $f \in L_\infty$ and every probability density p with respect to μ, where "$*_\mu$" denotes convolution with respect to μ,

$$p *_\mu f(s) = \int p(g) f(g^{-1}s)\, d\mu(g).)$$

(ii) By definition, a topologically right invariant mean m on L_∞ satisfies $m(f*p)=m(f)$ for all $f\in L_\infty$ and all probability densities p. Such a mean is in fact invariant in the following sense, which is stronger than being invariant in the sense of either M or J; namely, $m(f*\pi)=m(f)$ for all $f\in L_\infty$ and all regular Borel probability measures π on G (where

$$f*\pi(g)=\int f(gs^{-1})\,d\pi(s)$$

a.e. (v) (Hewitt and Ross (1963), (20.12) Theorem (i)). This follows from Hewitt and Ross (1963), (19.18) Theorem.

(iii) Using the ideas of Remark (ii), one can easily show that every topologically right invariant mean on L_∞ (or UC) is right invariant. From this fact and Remark 2 at the beginning of the appendix, one can conclude that, if one of the subspaces considered here has a mean invariant in one of the senses under consideration, then L_∞ has a mean simultaneously invariant in all those senses.

(iv) Every right invariant mean on RUC is topologically right invariant (Greenleaf (1969), proof of Lemma 2.2.2). However on $C(G)$ and L_∞ there can exist right invariant means that are not topologically right invariant (Rosenblatt (1976, 1978)). For example, on $L^\infty(T)$, where T is the circle group, there exist 2^c "mutually singular" (right) invariant means, only one of which can be topologically invariant, since $C(T)$ has a unique invariant mean.

Condition WC (Day's weak convergence to invariance) – there exists a net $\{h_\alpha\}$ of probability density functions such that, for every $g\in G$, $h_\alpha-(h_\alpha)_g\to 0$ weakly in $L_1(G)$ (i.e.,

$$\int_G [h_\alpha(s)-h_\alpha(sg)]\,f(s)\,dv(s)\to 0$$

for every $f\in L_\infty(G)$). It follows from basic Banach space theory that WC is equivalent to M.

Condition WCT (the "topological" version of WC) – there exists a net $\{h_\alpha\}$ of probability density functions such that, for every probability density function h, $h_\alpha-h_\alpha*h\to 0$ weakly in $L_1(G)$ (i.e., for every $f\in L_\infty(G)$,

$$\int [h_\alpha(g)-(h_\alpha*h)(g)]f(g)\,dv(g)\to 0).$$

Banach space theory yields directly the equivalence of this condition and J. See the proof $J\Rightarrow$ WCT in the Appendix.

Condition SCT (strong convergence to topological invariance) – there exists a net $\{h_\alpha\}$ of probability density functions such that, for every probability density function h, $h_\alpha-h_\alpha*h\to 0$ in $L_1(G)$ (i.e.,

$$\|h_\alpha-h_\alpha*h\|_1=\int |h_\alpha(g)-(h_\alpha*h)(g)|\,dv(g)\to 0).$$

Condition P – for each compact $K\subset G$ and $\varepsilon>0$, there is a probability density function h such that, for every $g\in K$,

$$\|h_g-h\|_1=\int |h(g'g)-h(g')|\,dv(g')<\varepsilon.$$

This is H. Reiter's condition, which he called (P_1) and applied in a number of directions (see Reiter (1968), Chap. 8, or Greenleaf (1969), Sect. 3.2, for details).

In the following definition, and in all others where "$\nu(G_\alpha)$" occurs in a denominator, it is understood that $\nu(G_\alpha)>0$.

Condition S (existence of a summing net) – there is a net $\{G_\alpha\}$ of compact subsets of G with $G_\alpha \supset G_\beta$ whenever $\alpha \geqq \beta$ and $\bigcup_\alpha G_\alpha = G$ and such that $\nu(G_\alpha g \cap G_\alpha)/\nu(G_\alpha) \to 1$ uniformly (in g) on compact subsets of G. Whenever such a net exists, it can be chosen so that

(i) the G_α's are symmetric ($G_\alpha = G_\alpha^{-1}$).

In case G is σ-compact, the net $\{G_\alpha\}$ can be chosen to be a sequence (summing sequence), in accordance with the earlier general remarks about σ-compact G. And, when G is unimodular, the G_α's can be chosen to satisfy both (i) and also

(ii) $\nu(G_\alpha K \cap G_\alpha)/\nu(G_\alpha) \to 1$ for each compact $K \subset G$.

For non-unimodular amenable G, it is not known if there is a summing net simultaneously satisfying both (i) and (ii); it seems reasonable to conjecture that there is such a summing net. It *is* known that every locally compact amenable group, unimodular or not, admits a summing net satisfying (ii), but perhaps not (i). (See Emerson (1974a) in this regard.)

Condition A_1 – there exists a net $\{G_\alpha\}$ of compact sets such that $\nu(G_\alpha K)/\nu(G_\alpha) \to 1$ for all compact $K \subset G$. (Note that the left-handed form has $\mu(KG_\alpha)$ in the numerator, *not* $\mu(G_\alpha K)$; in fact the condition $\mu(G_\alpha K)/\mu(G_\alpha) \to 1$ cannot be satisfied, unless G is unimodular, even if we restrict ourselves to the singleton sets $K=\{g\}$, $g \in G$.)

Condition FW_1 – for each finite $F \subset G$ containing the identity e, there exists a sequence $\{G_n\}$ of Borel sets with $\nu(G_n)<\infty$ such that

$$\nu\Big(\bigcap_{g \in F} G_n g\Big)/\nu(G_n) \to 1.$$

Condition GR – the constant function **1** (equal everywhere to 1) can be approximated uniformly on every compact subset of G by continuous positive definite functions vanishing outside compact sets. This condition was used by Grenander ((1963), Chap. 5) to prove probabilistic limit theorems on groups using Fourier transforms. An astounding point about this condition is that it is expressed entirely in terms of functions that are constant or vanish at infinity; and the algebra generated by such functions always has an invariant mean, whether G is amenable or not. The equivalence of this condition and P was shown essentially by H. Reiter (1964); see also Reiter (1968), Sect. 8.3. (An excellent overview of the role and history of positive definite functions in mathematics has been written by Stewart (1976).)

Condition SC (strong convergence to invariance) – there exists a net $\{h_\alpha\}$ of probability density functions such that

$$\|(h_\alpha)_g - h_\alpha\|_1 = \int |h_\alpha(g'g) - h_\alpha(g')|\, d\nu(g') \to 0 \qquad \text{for all } g \in G.$$

(Stone and von Randow (1968) call this condition P_2.)

Condition P_3 – For each pair $\{g_1, g_2\} \subset G$, there exists a sequence $\{h_n\}$ of probability density functions such that

$$\|(h_n)_{g_i} - h_n\|_1 \to 0, \qquad i = 1, 2.$$

The analogous condition with "pair $\{g_1, g_2\}$" replaced by "finite set of points" is easily seen to imply amenability (and is, in fact, equivalent to it). For some implications of P_3 in statistical theory, see Stone and von Randow (1968), Sect. 4.

Condition HS – there is a net $\{P_\alpha\}$ of probability measures defined on the Borel sets of G such that for all $g \in G$ and Borel $B \subset G$, $|P_\alpha(Bg) - P_\alpha(B)| \to 0$. (The statement of the left-handed version uses "$P_\alpha(gB)$" and is otherwise identical.) HS is the condition of the Hunt-Stein theorem (Lehmann (1959), p. 336). It is usually stated in terms of sequences rather than nets; in effect, this restricts the theorem to σ-compact groups.

Torgersen (1972), p. 1387, asserts the equivalence of HS and amenability; we give a sketch of the proof in Sect. 3. B. Zehnwirth pointed out to us that amenability is also equivalent to the stronger condition (to be called HSU), where, for each $g \in G$, the convergence in HS is uniform in B.

Condition H_1 – there exists a net $\{G_\alpha\}$ of closed sets with $\nu(G_\alpha) < \infty$ such that

$$\nu\Big(\bigcap_{k \in K} G_\alpha k\Big) / \nu(G_\alpha) \to 1$$

for every compact K. This condition was shown by Bondar (1977) to imply that "strong inconsistency" cannot occur between Neyman-Pearson and flat Bayes confidence intervals.

Condition H_2 – for each pair $\{g_1, g_2\} \subset G$, there exists a sequence $\{G_n\}$ of Borel sets with $\nu(G_n) < \infty$ for all n such that

$$\nu(G_n \cap G_n g_1 \cap G_n g_2) / \nu(G_n) \to 1.$$

Condition Π_1 – there exists a net $\{G_\alpha\}$ of Borel sets with $\nu(G_\alpha) < \infty$, such that $\nu(G_\alpha \cap G_\alpha g)/\nu(G_\alpha) \to 1$ for all $g \in G$. This is the right-handed form of the "weak boundedness" of Peisakoff's 1950 thesis on minimaxity.

Solvability – there is a finite chain

$$G = G_m \supset G_{m-1} \supset \ldots \supset G_0 = \{e\}$$

of subgroups terminating in the identity, such that G_j is the commutator subgroup (the closure of the subgroup generated by all elements of the form $g_1 g_2 g_1^{-1} g_2^{-1}$) of G_{j+1}, $0 \leqq j < m$. This implies that each G_{j+1}/G_j is a commutative group. In the context of Lie group theory, the term "solvable" is used to refer to a similar property defined in terms of chains of subalgebras of the Lie algebra of G. The above definition, however, is the one occurring in statistics, e.g., Stein (1965). Solvability implies

Stein's Condition – there is a finite chain

$$G = G_m \supset G_{m-1} \supset \ldots \supset G_0 = \{e\}$$

of closed subgroups, each a normal subgroup of the previous one, such that each G_i/G_{i-1} is either compact or commutative. This condition appears in Stein (1965), p. 225. If G satisfies Stein's condition, then such a chain can be found for which G_i/G_{i-1} is compact if $i=m$ and commutative if $i<m$. Iwasawa (1949), in his work on Hilbert's fifth problem, was led to consider connected groups satisfying Stein's condition; he called such groups (C)-groups.

G has the *fixed-point property* if every representation $g \to T_g$ ($T_g(x)$ being separately continuous in g and x) of G as a group of affine transformations of a compact convex subset K of a locally convex topological vector space has a fixed point (i.e., there exists $k \in K$ such that $T_g(k)=k$ for all $g \in G$). For abelian groups, the fixed-point property is established by the Markov-Kakutani fixed-point theorem (Markov (1936), Kakutani (1938), or Dunford and Schwartz (1958), Theorem V.10.6). Furstenberg (1963) proved that, for connected Lie groups, the fixed-point property and Stein's condition are equivalent. Proceeding from the work of Day (1961), Rickert (1967) proved the equivalence, in general, of these two conditions. (See also Greenleaf (1969), Sect. 3.3.)

Condition K – whenever G leaves invariant the statistical problem of testing the G-invariant hypotheses H_0 against H_1, and certain regularity conditions are satisfied (namely, G is locally compact and σ-compact, acting measurably on the second countable sample space $\mathscr{X}$, and the probabilities P_θ, $\theta \in H_0 \cup H_1$, on $\mathscr{X}$ are dominated by a σ-finite measure m), *then*, for every randomised test function ψ of size α defined on $\mathscr{X}$, there is an invariant randomised test function $\bar{\psi}$ also of size α, such that

$$\inf_{g \in G} E_{g\theta} \psi(X) \leqq E_\theta \bar{\psi}(X) \leqq \sup_{g \in G} E_{g\theta} \psi(X), \qquad \theta \in H_0 \cup H_1.$$

Here, *test function* means any $\psi \in L_\infty(\mathscr{X}, m)$ for which $0 \leqq \psi(x) \leqq 1$, $x \in \mathscr{X}$. One form of the Hunt-Stein theorem is the statement "HS $\Rightarrow$ K". The reverse implication K $\Rightarrow$ HS is also true for almost connected groups (Bondar and Milnes (1977)), giving a converse to the Hunt-Stein theorem.

The Emerson-Templeman Condition – there exists a summing net $\{G_\alpha\}$ of compact sets for which

(*) there is an upper bound $B < \infty$ for the ratios $\nu(G_\alpha G_\alpha^{-1})/\nu(G_\alpha)$. This condition was formulated by Templeman (1967) in an announcement of a pointwise (individual) ergodic theorem for groups of transformations; see Emerson (1974) for statement and proof. It has been suggested that the existence of a net $\{G_\alpha\}$ of compact subsets of G satisfying $\bigcup_\alpha G_\alpha = G$ and (*) ought to imply amenability. We do not know if this is the case.

It will be noted that some of our statements involving intersections can be reformulated using the symmetric difference $A \Delta B$ $(=(A \cup B)-(A \cap B))$. For example, condition S requires that $\nu(G_\alpha \Delta G_\alpha g)/\nu(G_\alpha) \to 0$. Some other statements involving intersections can be reformulated in a different way, since, for any given $a>1$ and compact $K\subset G$, the assertion that there is a (measurable) set U with

$$\nu(\bigcap_{g\in K} Ug)/\nu(U)\geqq 1/a$$

is equivalent to the assertion that there is a set U' with

$$\nu(U'K)/\nu(U')=\nu(\bigcup_{g\in K} U'g)/\nu(U')\leqq a.$$

Thus H_1 would be equivalent to A_1 if the net $\{G_\alpha\}$ in H_1 was required to consist of compact sets. (The reader can write down for himself the precise "union" equivalents of conditions H_1, H_2 and Π_1.)

We may also note that many of the conditions occur in pairs, one condition involving a ratio of measures of sets, the other involving probability measures; for example, the pair H_2 and P_3, and the pair Π_1 and SC. To be more precise, if $\{G_n\}$ is a sequence satisfying H_2 and we normalise the indicator functions of the G_n's, we get a sequence of probability densities $\{h_n\} =\{\nu(G_n)^{-1}I_{G_n}\}$, which by an easy calculation satisfies P_3. Thus, moving from a statement about ratios of measures of sets to one about probability densities is straightforward. On p. 337 Lehmann (1959) considers strengthening condition HS by requiring the probability measures in HS to be of the form $P_n =\nu(G_n)^{-1}I_{G_n}$ for suitable $G_n\subset G$. One easily sees that condition S implies this strengthened HS condition.

In addition to the above, many more conditions may be found in Stone and von Randow (1968), Greenleaf (1969), Day (1969), and other recent works on functional analysis such as Berg and Christensen (1974), Chou (1980), Emerson (1978, 1979), Eymard (1975), Furstenberg (1963), Gilbert (1968), Glasner (1976), Guivarc'h (1973), Herz (1973), Jenkins (1973), Johnson (1972, 1977), Lance (1973), Milnes (1978), Paschke (1978), Rindler (1976), Schwartz (1963) and Sherman (1979).

Theorem 1.1. *All the conditions defined above are equivalent for locally compact Hausdorff groups G, save for the Emerson-Templeman condition, Stein's condition and solvability, each of which implies amenability, and K, H_2 and P_3 which are implied by it. If G is almost connected, then K, P_3, H_2 and Stein's condition are each equivalent to amenability. Each of K, P_3 and H_2 is equivalent to amenability if G is a subgroup of a linear group furnished with the discrete topology.*

G is *almost connected* if G/G_0 is compact, where G_0 is the (connected) component of the identity. In particular, any group which is connected or has a finite number of connected components is almost connected. By a *linear group*, we mean the group $GL(V)$ of invertible linear transformations of a finite dimensional real vector space V. It is worth noting that any group of affine transformations of a finite dimensional vector space can be embedded as a subgroup of a linear group.

A diagram of the implications of the theorem:

$$\begin{array}{ccccccc}
 & & \text{Solvable} & & & & \\
 & & \Downarrow & & & & \\
 & & \text{Stein} & & & & \\
 & & \Downarrow\uparrow & & \Lleftarrow & & \\
\text{Emerson-Templeman} & \Rightarrow & \text{Amenability} & \Rightarrow & H_2 & \Leftrightarrow & P_3 \\
 & & \uparrow\Downarrow\Uuparrow & & \leftarrow & & \\
 & & K & & & &
\end{array}$$

($\rightarrow$ indicates implication with the added hypothesis that G be almost connected; $\Rrightarrow$ indicates implication with the added hypothesis that G be a subgroup of a linear group.) Proofs of the various parts of Theorem 1.1 are discussed in Sect. 3 and the appendix.

It is well known that solvability is not necessary for amenability; for example, the symmetric group on 5 symbols is not solvable (a fact proved in Galois theory to be equivalent to the impossibility of solving all quintic equations in radicals - hence the name), but this group is finite, hence amenable. It is not known if the Emerson-Templeman condition holds for any amenable non-unimodular groups, or even for all amenable unimodular groups.

There is a celebrated conjecture of von Neumann (1929) that every non-amenable group contains a free group on two generators as a closed subgroup. If this is true, the free group on two generators is in a sense the basic non-amenable group. Stone and von Randow have remarked that a P_3 group cannot contain a free group on two generators as a closed subgroup. Therefore, if von Neumann's conjecture is true, then a group which is not amenable cannot satisfy P_3, hence $P_3 \Rightarrow$ amenable (without further hypotheses on the group in question). The conjecture is known to be true for almost connected groups (Rickert (1967), Theorem 5.5) and for subgroups of linear groups (Tits (1972)), the latter furnished with the discrete topology. Fortunately, perhaps, all the groups currently used in parametric statistics (save for the infinite discrete groups) have a finite number of connected components and *a forteriori* are almost connected. The word "closed" in the statement of the conjecture is critical; for, the rotation group $SO(3)$, which is compact and thus amenable, contains the free group on two generators as a (non-closed) subgroup; von Neumann used this fact to resolve the Hausdorff-Banach-Tarski "paradox" (Greenleaf (1969), Sect. 1.3). Sherman (1979) and Emerson (1979) independently have formulated an analogue of this paradox for locally compact groups so as to provide another characterization of amenability.

2. The Class of Amenable Groups

The following are basic and well known (Greenleaf (1969), Day (1969)); for discrete groups Theorems 2.3–2.6 were proved by von Neumann (1929), who initiated the study of invariant means.

Theorem 2.1. *All compact groups are amenable. (The proof is easy: if ν is normalised Haar measure, $\nu(G)=1$, then $f \to \int_G f(g)\,d\nu(g)$ is an invariant mean.)*

Theorem 2.2. *All abelian groups are amenable.*

Theorem 2.3. *Any closed subgroup of an amenable group is amenable.*

Theorem 2.4. *The image of an amenable group under a continuous homomorphism is amenable.*

Theorem 2.5. *If N is a closed normal subgroup of G and if N and G/N are amenable, then G is amenable.*

Corollary 1. *The direct product of two groups is amenable if and only if both of these groups are amenable. (The "only if" part follows from 2.3; the "if" part follows from 2.5.)*

Corollary 2. *The semidirect product of two amenable groups is amenable.*

Theorem 2.6. *If G is a directed union of a system of closed amenable subgroups $\{H_\alpha\}$, in the sense that $G=\bigcup_\alpha H_\alpha$ and for any H_α, H_β there exists H_γ such that $H_\gamma \supset H_\alpha \cup H_\beta$, then G is amenable.*

Some amenable groups are: the usual additive real numbers ($\{G_n\}=\{[-n,n]\}$ is a summing sequence as shown in Example 1), the positive real numbers with ordinary multiplication $\left(\{G_n\}=\left\{\left[\frac{1}{n},n\right]\right\}\text{ is a summing sequence}\right)$, the translations of a finite-dimensional vector space, the scalar multiplications on a vector space (the "scale group"), the "translation-scale group" of the real line (also known as $GA(1)$), the group of the Behrens-Fisher problem (which is the direct product of two translation-scale groups, hence amenable by Corollary 1), the group $T(n)$ of non-singular upper triangular matrices (the "triangular group"; this group is even solvable), the group generated by $T(n)$ and the translations in E^n (since this is a semi-direct product of the triangular and translation groups). Groups satisfying Stein's condition are amenable (since G_1 in the chain of normal groups defining Stein's condition, is commutative or compact, hence amenable. Now induction on m using 2.5 shows that G is amenable); solvable groups are a special case. The univariate two-sample problem is invariant under the group of transformations of the form

$$(x_1,x_2,\ldots,y_1,y_2,\ldots)\to(sx_1+a,sx_2+a,\ldots,sy_1+b,sy_2+b,\ldots),\qquad s\neq 0;$$

this group is amenable. Consider the group S_∞ consisting of those permutations of a countable collection of objects which leave all but a finite number of the objects fixed; this group is of interest as the largest group leaving invariant all exchangeable probabilities. S_∞ is amenable by Theorem 2.6, and its invariant means were used as priors for a Bayesian sampling model in Lane and Sudderth (1978).

Some groups which are *not* amenable: the free group on two generators (which was used in 1950 by Peisakoff in a counterexample to a conjecture regarding minimax decision rules), the general linear group $GL(n)$ for $n \geqq 2$ (i.e., the non-singular $n \times n$ matrices; that this group fails to satisfy HS was shown by Stein (Lehmann (1959), Sect. 8.4, Example 9)), the group generated by translations and non-singular matrices in E^n for $n \geqq 2$ (known as the general affine group $GA(n)$). The group $SL(n)$ of $n \times n$ matrices with unit determinant is not amenable for $n \geqq 2$ (which follows from 2.5 since $GL(n)/SL(n)$ is the group of reals).

It follows from 2.3 and 2.5 that, if G has a finite number of connected components, then G is amenable iff its connected component of the identity is amenable. In particular, if such a group G is a closed subgroup of $GL(n)$, then G is amenable iff the subgroup consisting of those elements of G with positive determinant is amenable. The usual full group leaving invariant the hypothesis $\boldsymbol{\mu}=\mathbf{0}$ in MANOVA, namely $GL(n) \times O(r) \times E^{nr}$ (Lehmann (1959), Chap. 7.10) is not amenable for $n \geqq 2$, but the group (isomorphic to $O(n) \times O(r) \times E^{nr}$) leaving invariant the hypothesis $\Sigma = \Sigma_0$ is amenable.

Remarks on Applications. These remarks supplement comments in the last three paragraphs and following the definitions of conditions GR, P_3, HS, H_1, Π_1 and K.

Probably the first statistical application of amenability was the testing version of the Hunt-Stein theorem, c. 1946. The best presentation of this in the literature is in Lehmann (1959), Chap. 8, Theorem 2. We shall shortly give a more direct proof using the fixed point property.

In the following, a *G-invariant statistical model* $(\mathscr{X}, \mathscr{B}, m, \{P_\theta | \theta \in \Omega\})$ consists of a positive measure space $(\mathscr{X}, \mathscr{B}, m)$ for which $L_1(\mathscr{X}, \mathscr{B}, m)^* = L_\infty(\mathscr{X}, \mathscr{B}, m)$ and a set Ω indexing a set of probability densities P_θ in $L_1(\mathscr{X}, \mathscr{B}, m)$ (i.e., each P_θ is a probability density with respect to m on $\mathscr{X}$). (See Dunford and Schwartz (1958), pp. 289–90, concerning the requirement $L_1^* = L_\infty$.) G acts as a set of transformations of $\mathscr{X}$, i.e., a function

$$(g, x) \to gx, \qquad G \times \mathscr{X} \to \mathscr{X}$$

is defined satisfying $g_1(g_2 x) = (g_1 g_2)x$ for all $x \in \mathscr{X}$ and $g_1, g_2 \in G$. We assume that $\mathscr{B}$ is G-invariant, i.e., $gB = \{gx | x \in B\} \in \mathscr{B}$ for all $B \in \mathscr{B}$, and that, if gm is defined for $g \in G$ and $B \in \mathscr{B}$ by $gm(B) = m(gB)$, then m dominates each such gm. It follows for $g \in G$ that the map

$$f \to gf, \qquad gf(x) = f(g^{-1}x)$$

is an isometry of L_∞ onto L_∞ and its adjoint

$$h \to g^{-1}h, \qquad g^{-1}h(x) = h(gx)\frac{d(gm)}{dm}(x)$$

is an isometry of L_1 onto L_1. (Note $(gf, h) = (f, g^{-1}h)$.) We assume as well that, for each $h \in L_1$, the map

$$g \to gh, \qquad G \to L_1$$

is norm continuous and that $\{P_\theta | \theta \in \Omega\}$ is G-invariant, i.e.,

$$g P_{\theta'} = P_{g\theta'} \in \{P_\theta | \theta \in \Omega\}$$

for all $\theta' \in \Omega$.

These hypotheses for a G-invariant model allow a proof of the Hunt-Stein theorem requiring hardly more than an application of the fixed point property. After the theorem we will indicate why these hypotheses are not as restrictive as they might appear. For terms and notation used in the theorem and not defined here, the reader is referred to Lehmann (1959).

Theorem 2.7 (Hunt-Stein). *Let G be a locally compact group with the fixed point property and let $(\mathscr{X}, \mathscr{B}, m, \{P_\theta | \theta \in \Omega\})$ be a G-invariant statistical model. Then, for every critical function ψ of size α and power function $\beta(\theta)$ for testing an invariant hypothesis H_0 against an invariant hypothesis H_1, there exists an almost invariant critical function ψ' of size at most α and power function $\beta'(\theta)$ such that*

$$\inf_{g \in G} \beta(g\theta) \leqq \inf_{g \in G} \beta'(g\theta), \qquad \theta \in H_1.$$

Proof. Let $C_{\alpha\beta}$ be the subset of the unit ball of L_∞ consisting of those f's for which $E_\theta(f) \leqq \alpha$ for all $\theta \in H_0$ and $E_\theta(f) \geqq \inf_{g \in G} \beta(g\theta)$ for all $\theta \in H_1$. $C_{\alpha\beta}$ is non-void, since the function class in L_∞ determined by ψ is in it, and $C_{\alpha\beta}$ is a weak* closed subset of the unit ball of L_∞, hence is weak* compact. $C_{\alpha\beta}$ is also convex; and the action of G on L_∞ is by linear (hence affine) maps, leaves $C_{\alpha\beta}$ invariant and is separately continuous. (For example, if $f \in L_\infty$, one must show that the function

$$g \to gf, \quad G \to L_\infty$$

is weak* continuous, i.e., that the function

$$g \to (gf, h) = \int f(gx)\, h(x)\, dm(x)$$

is continuous for each $h \in L_1$. But this follows immediately from the norm continuity of the map $g \to gh$ and the fact that $(gf, h) = (f, g^{-1}h)$. The continuity of the map

$$f \to gf, \quad C_{\alpha\beta} \to C_{\alpha\beta}$$

is easier to establish.) Hence, the fixed point property yields an $f' \in C_{\alpha\beta}$ such that $gf' = f'$ for all $g \in G$. Any member ψ' of the function class f' then has the desired properties.

We wish to make one point immediately: it follows from Theorem 4 in Chap. 6 of Lehmann (1959) that one can often get an invariant test function from an almost invariant one.

Generality of the Hypotheses. We would like to draw attention to some results mainly of functional analysis.

1. In our definition of G-invariant model, we assumed the map $g \to gf$, $G \to L_1$ was continuous for each $f \in L_1$. This conclusion often follows from the

assumption that these maps are merely weakly measurable; see Moore (1968), Chap. 4.

2. By Theorem H, p. 275, of Halmos (1950), a measurable group G_1 can often be viewed as a special, in particular, dense, subgroup of a locally compact group G. It may then be possible to extend a weakly measurable representation of G_1 to a weakly measurable representation of G.

3. In this discussion of hypotheses for the Hunt-Stein theorem, it seems appropriate to remind the reader of Lemma III.8.5 in Dunford and Schwartz (1958) and of the example on pp. 153–4 in de Leeuw and Glicksberg (1965), and also to point out that $L_1(\mathscr{X}, \mathscr{B}, m)$ need not be separable even when m is finite.

Lehmann (1959a) has generalized the Hunt-Stein theorem by considering functions more general than $\inf_{g\in G} \beta(g\theta)$, such as the regret function which gives the existence of most stringent tests. Further, the Hunt-Stein theorem has been extended (Kiefer (1957), Wesler (1959)) to estimation and other decision spaces: under fairly general topological conditions, if G is amenable and any decision function δ_0 is given for an invariant statistical decision problem, then an equivariant decision function δ' (i.e., one for which $\delta'(gx|gA)=\delta'(x|A)$) exists such that the risks R satisfy

$$\sup_{g\in G} R_{\delta'}(g\theta) \leqq \sup_{g\in G} R_{\delta_0}(g\theta), \qquad \theta\in\Omega. \tag{1}$$

It is this type of result which we shall call the generalized Hunt-Stein theorem. The proofs of it in the literature are technical and require extraneous conditions to circumvent measure-theoretic difficulties. Cleaner and more elegant proofs can be given using the fixed point property, an idea first used by Le Cam and later by Huber. Regrettably, none of these proofs (Brown (1980), Portnoy (1975), Zehnwirth (1977)) have yet been published, although a brief outline of such a proof is in Kiefer (1966), p. 263.

Here is an outline of a proof using the same approach: let δ_0 be a decision function, i.e., $\delta_0(x, D)$ is the probability that if x is observed then a decision will be chosen from the subset D of the space $\mathscr{A}$ of all possible decisions. Then

$$(T_g\delta)(x, D) = \delta(gx, gD)$$

for all g, δ, x and D defines a representation T of G by affine maps T_g. Let S be the set of δ' satisfying inequality (1) above; S is G-invariant and convex and, under fairly general conditions, will be compact in a suitable locally convex topology, for which the action of G on S is separately continuous. Now, by the fixed point property, there is a $\delta'\in S$ for which $\delta'(gx, gD)=\delta'(x, D)$ for every $g\in G$ and $D\subset\mathscr{A}$. Since decision functions are identified if they agree for almost all x, this is the condition of almost equivariance of δ'.

The difficult part of this proof is to find the suitable locally convex topology on the class of decision functions. To get from an almost equivariant δ' to an equivariant δ' we must invoke an equivariant version (Berk and Bickel (1968), see also comments in Berk (1970)) of Theorem 4 in Chap. 6 of

Lehmann (1959). It says that δ' is almost equivariant iff δ' is almost everywhere equal to an equivariant decision function – provided one has mild measurability conditions, plus the condition that there exist at least one equivariant δ, not necessarily in S. (An equivariant δ need not exist; see Berk (1967).)

It is to be noted that if G acts freely and transitively on both the parameter and sample spaces of an invariant statistical model, and if $f(x)$ is the density function with respect to μ of the observation when $\theta=e$, then the power function $\beta(\theta)$ of a test function $\psi(x)$ equals the convolution

$$f *_\mu \check{\psi}(\theta^{-1}) = \int \psi(x) f(\theta^{-1}x)\, d\mu(x).$$

This equality explains why some of our conditions on groups which involve convolutions are of statistical interest in their left-handed versions. If G is not transitive on $\mathscr{X}$, which is usually the case, then, typically, the power function is obtainable by integrating such convolutions with respect to the marginal probability of the maximal invariant statistic (Bondar (1976)).

Amenability of G is necessary in the Hunt-Stein theorems: Portnoy (1975) constructs a counterexample to the generalised Hunt-Stein theorem for any non-amenable G. In Bondar and Milnes (1977) using an idea of Peisakoff, a counterexample to the (ungeneralized) Hunt-Stein theorem is constructed for F_2, the free group on two generators: a critical function ψ and two probability measures P_1 and P_2 on F_2 are exhibited which satisfy

$$P_1 *_\mu \check{\psi}(g) \geqq 1/2 \quad \text{and} \quad P_2 *_\mu \check{\psi}(g) \leqq 1/8$$

for all $g \in F_2$. It follows that ψ is a critical function of size $\alpha \leqq 1/8$ and power $\geqq 1/2$ for testing $\{(P_1)_g | g \in F_2\}$ against $\{(P_2)_g | g \in F_2\}$; and there can be no invariant ψ' with $\alpha' \leqq 1/8$ and power $\geqq 1/2$, since in this setting an invariant ψ' with $\alpha' \leqq 1/8$ also has power $\leqq 1/8$. Bondar and Milnes also "extend" this counterexample to any locally compact group G containing F_2 as a closed subgroup (e.g., to non-amenable almost connected G).

For the next four paragraphs, we assume that G acts exactly and transitively on the set Ω of an invariant statistical model, and consider the estimation of θ; our statements of hypotheses will not be as precise as in the previous paragraphs. Take a sequence of prior probability density functions (w.r.t. ν) on G,

$$\{\pi_n\} = \{\nu(G_n)^{-1} I_{G_n}\},$$

where $\{G_n\}$ is an increasing sequence of sets whose union is G. Then for each $x \in \mathscr{X}$, the sequence of posterior distributions $\{B_n(\cdot|x)\}$ converges weakly to $B_\nu(\cdot|x)$, the posterior induced from the prior measure ν. However, the convergence is not uniform in x, and in general each B_n differs greatly from B_ν on a set of x of high probability, even for large n. This leads us to examine the convergence more closely: let

$$d(x) = \sup_{A \subset G} |B_n(A|x) - B_\nu(A|x)|,$$

and consider X_n to be the random variable with the marginal distribution induced from the Bayesian joint distribution which gives θ the probability density function π_n and X the distribution P_θ; now $d(X_n)$ is a random variable. Stone (1970) shows, under regularity conditions, that $\{d(X_n)\}$ converges to zero in probability iff G is amenable and $\{G_n\}$ is chosen properly – a summing sequence will do.

This work of Stone sheds light on the claim often made in Bayesian circles that the Bayes posterior B_v can be viewed as an approximation to B_n for large n. (Since B_n arises from a proper prior, it has coherence and other pleasant properties which one could then hope to be at least partially shared by B_v.) This claim is true if G is amenable, and not otherwise in general.

Here are some properties of the B_n's that are (partially) shared by B_v. Heath and Sudderth (1978) show B_v to possess a coherence under betting somewhat weaker than the coherence defined in terms of Bayes betting possessed by the B_n's. Related to this is the consistency between B_v and Neyman-Pearson confidence regions: Stone (1976) exhibits an amusing invariant model with $G=F_2$ and a set estimator for θ whose Bayes credibility level (i.e., the probability of covering θ given x as evaluated by B_v) is $\gamma(x)=75\,\%$ for all x, yet, as a Neyman-Pearson confidence set estimator, it has a confidence level of only $\alpha(\theta)=25\,\%$ for all θ. Such strong inconsistency between the Bayes and Neyman-Pearson points of view cannot occur if G is amenable (Bondar (1977)); in fact, if m is an invariant mean on G and we define $f_x(g)=\gamma(gx)$ and $h(g)=\alpha(gx)$, then, under regularity conditions, $m(f_x)=m(h)$ for all x (Heath and Sudderth (1978)). If δ_n (resp. δ_v) is the Bayes decision rule relative to π_n (resp. v) for some fixed loss function, and if r_n (resp. r_{vn}) is the π_n-Bayes risk of δ_n (resp. δ_v) then $r_n - r_{vn} \to 0$ as $n \to \infty$ provided G is amenable, and the loss function is bounded; furthermore, δ_v is ε-admissible for all $\varepsilon > 0$. One might also ask under what circumstances $\{\delta_n\}$ converges to δ_v in some sense, but the answer appears to be complex. Stein (1965) looks at the convergence of $\{B_n\}$ to B_v in terms of information for G obeying Stein's condition.

Heath and Sudderth (1978) also show that any invariant mean m on G may be used as a prior probability (not σ-additive) and gives the same posterior as the measure v. The invariant mean m is more pleasing than v from a Bayes point of view since m gives measure 1 to Ω, unlike v which typically gives $v(\Omega)=\infty$. The invariance of m makes it truly non-informative. Zehnwirth (1975) shows that a least favourable sequence of prior distributions for estimating θ may be chosen to satisfy the asymptotic invariance property of condition P.

Le Cam (1964), Sect. 4, noted a connection between invariant means and the comparison of invariant experiments; this has been followed up by Torgersen (1972), see also Le Cam (1974), Sect. 3.

It has been shown (Kesten (1959)) that if G is a countable group, then the spectral radius of the transition matrix of a symmetric random walk (whose support generates G) is equal to one iff G is amenable.

Some of the considerations in this paper generalize to semigroups, see Day (1957, 1968, 1969) and Lau (1972). In particular, Theorem 2.7 holds if G is one of a class of semigroups.

Knowledge concerning the existence of invariant means and asymptotically invariant sequences of probabilities defined on a space acted on by a group of transformations is in a more primitive state than the theory for such means and probabilities on the group itself. If H is a closed subgroup of G, then there is an invariant mean on the quotient space G/H if G is amenable (Greenleaf (1969a)) but the converse is false. A survey of the field was written by Eymard (1972).

Ergodic theorems for amenable G acting measurably on a space $\mathscr{X}$ are found in Templeman (1967), Emerson (1974), Greenleaf (1973), Greenleaf and Emerson (1974). A typical result is as follows: if $f \in L_p(\mathscr{X}, \mathscr{B}, \lambda)$, $1 < p < \infty$, where λ is an invariant σ-additive measure and

$$A_n f(x) = (\nu(G_n))^{-1} \int_{G_n} f(xg)\, d\nu(g),$$

where $\{G_n\}$ is a summing sequence, then the sequence $A_n f$ converges in L_p norm to an invariant limit f^*. If the summing sequence also satisfies the Emerson-Templeman condition, then $\{A_n f\}$ converges a.e. (ν) to f^* (although Emerson has also shown that the Emerson-Templeman condition is not necessary to get this last conclusion). If G is the integers and $G_n = \{j \in G \mid -n \leqq j \leqq n\}$, then these facts are the classical L_p and individual ergodic theorems, respectively. To be found in Eberlein (1949) is a complete treatment of mean ergodic theory, done in the setting of bounded linear operators on a Banach space.

The reader must be warned that not every individual ergodic "theorem" in the literature is true (see the comments in Emerson (1974b)). We remark that, although the sequence $\{[-n, n]\}$ of subsets of R satisfies the Emerson-Templeman condition, the sequence $\{A_n F\} \subset L_\infty$ defined by

$$A_n F(x) = (2n)^{-1} \int_{-n}^{n} F(x+y)\, dy, \qquad x \in R$$

where F is defined in Example 1 of Sect. 1, diverges for *every* x. Thus there is no analogue for L_∞ of the pointwise ergodic theorem that holds for L_p spaces, $1 < p < \infty$.

Explicit construction of summing sequences is done in Greenleaf and Emerson (1974), Sects. 4-6, and in Greenleaf (1973), Sects. 5-6, for connected Lie groups.

3. Proofs

For each of condition GR and the fixed-point property we mentioned right after the definition in Sect. 1 where a proof of the equivalence of that condition and amenability could be found; also, it is clear that solvable groups satisfy Stein's condition and Theorems 2.2 and 2.5 show that groups satisfying Stein's condition are amenable. In an appendix we start with condition M_{ucb} and proceed through conditions M, J, WCT, SCT and P to S; condition A_1 essentially appears in this chain as well. It is clear that condition S implies condition M (and all the others in the chain) and, with a little effort, that conditions FW_1, SC, H_1, Π_1 and WC imply M and are implied by S. Also, groups satisfying the Emerson-Templeman condition are amenable; the converse is not known.

As to the remaining implications in Theorem 1, we note that condition S readily implies each of HS, HSU, P_3, H_2 and K. The proofs of the remaining implications we will discuss in a little more detail. Among these proofs, only the last (5) has appeared in the literature.

(1) *HS* $\Rightarrow$ *amenable.* This follows because the set $\mathscr{L}$ of linear combinations of characteristic functions is (norm-)dense in L_∞ and any weak* limit point m in L_∞^* of the net $\{P_\alpha\}$ is a right invariant mean on $\mathscr{L}$, hence on L_∞. (To be explicit, we note that HS implies directly that $m(f_g)-m(f)=0$ for all $f\in\mathscr{L}$ and $g\in G$; hence, if $F\in L_\infty$ and $f\in\mathscr{L}$ satisfies $\|F-f\|_\infty<\varepsilon$, then $\|F_g-f_g\|_\infty=\|F-f\|_\infty<\varepsilon$ and

$$|m(F_g)-m(F)|\leqq|m(F_g-f_g)|+|m(f_g)-m(f)|+|m(f-F)|\leqq 2\varepsilon.)$$

(2) $P_3 \Rightarrow H_2$. ($H_2 \Rightarrow P_3$ is trivial.) A proof can be conducted along the lines of (but is somewhat easier than) step 1 of the proof suggested in the appendix that $P \Rightarrow S$.

(3) P_3 *or* H_2 $\overset{?}{\Rightarrow}$ *amenable.* If G is P_3 (or, equivalently, H_2), then it cannot contain a closed subgroup isomorphic to the free group on two generators (Stone and von Randow (1968), Theorems 3.1 and 3.2); hence, it must be amenable if it is almost connected (Rickert (1967), Theorem 5.5) or if it is a subgroup of a linear group (Tits (1972)).

(4) K $\overset{?}{\Rightarrow}$ *amenable.* Citing Rickert and Tits as above, we conclude that G contains the free group on two generators as a closed subgroup if it is not amenable. The argument in Bondar and Milnes (1977) shows that such a group does not satisfy K.

(5) *Amenable* $\rightarrow$ *Stein.* (See Reiter (1968), 7.1.) If G is almost connected and amenable, then Theorem 5.3 of Rickert (1967) implies $G/\mathrm{rad}(G)$ is compact, where $\mathrm{rad}(G)$ is the radical of G, i.e., the (unique and closed) largest solvable connected normal subgroup of G. This completes the proof.

Remark. The authors would like to thank a referee for pointing out the relevance of the paper of Tits (1972) to our survey. Among other things, the referee pointed out the following analogue of (5) (see Tits (1972), Theorem 2). Let V be a vector space over a field k of characteristic different from 0 and let G be a subgroup of $GL(V)$ furnished with the discrete topology. If G is amenable, then G has a solvable normal subgroup H such that G/H is locally finite (i.e., a finite subset of G/H generates a finite subgroup of G/H. The amenability of locally finite groups follows from Theorem 2.6.)

Appendix

Our purpose here is to start with a not necessarily σ-compact, locally compact group G satisfying M_{ucb} and to give, all in one place, in a logical order, using right Haar measure, an account of the construction of a summing net for G as in condition S. (An alternate route for part of this construction appears in Eymard (1975). Also, limitation of space dictates that only an indication of proof can be given at most stages.) This net will be made to satisfy both (i) and (ii) of S if G is unimodular. If G is not unimodular, we can make the net satisfy either (i) or (ii) of S; it is not known if a summing net satisfying (i) and (ii) simultaneously exists in this case.

An account, as just described, has not yet appeared in the literature, although the various parts of it have, by and large, appeared (as will be mentioned in due course). The construction of summing nets for groups that are not σ-compact is, to our knowledge, not yet in the literature.

Remark 1. In the course of this appendix, it is shown that the summing net can be chosen to be a sequence if G is σ-compact; however, even in this case, the use of nets in the proof is unavoidable, and it can be shown only at quite a late stage in the proof that the summing net can be chosen to be a sequence.

As above, we will here identify the probabilty density functions (with respect to ν) with the functions $h \in L_1(G,\nu)=L_1$ that satisfy $\|h\|_1=1$ and $h \geqq 0$ a.e. (ν). We recall that the *modular function* Δ of G is defined by

$$\int f(sg)\,d\nu(g)=\Delta(s)\int f(g)\,d\nu(g), \qquad f\in L_1,\ s\in G$$

(which is equivalent to

$$\int f(gs)\,d\mu(g)=\Delta(s^{-1})\int f(g)\,d\mu(g), \qquad f\in L_1(G,\mu),\ s\in G$$

(Hewitt and Ross (1963), p. 195)), and that, if $h\in L_1$, then h^*, defined by

$$h^*(g)=h(g^{-1})\,\Delta(g), \qquad g\in G,$$

is in L_1 as well, $\|h^*\|_1=\|h\|_1$ and $(h^*)^*=h$; also h^* is a probability density if h is. G is called *unimodular* if $\Delta(g)=1$ for all $g\in G$.

Remark 2. If $\{h_\alpha\}$ is a net of probability density functions as in condition SCT, it follows that any weak* accumulation point in L_∞^* of $\{h_\alpha * h_\alpha^*\}$ is a topologically (left and right) invariant mean.

$M_{ucb} \Rightarrow M$ and J. Suppose m is a right invariant mean on LUC$\cap$RUC$=$UC, the bounded uniformly continuous functions on G. We note first that m is already topologically right invariant. (Two proofs of this fact appear in Greenleaf (1969).) If h is a probability density on G, one extends m to a topologically right invariant mean m_1 on RUC by the formula

$$m_1(f)=m(f*h), \qquad f\in \mathrm{RUC}.$$

(Recall such an $f*h$ is in UC.) m_1 is extended to a topologically right invariant mean m_2 on L_∞ by the formula $m_2(f)=m_1(\check{h}*f)$. (Recall such an $\check{h}*f$ is in RUC.) See Greenleaf (1969), Sect. 2.1, for details.

$J \Rightarrow WCT$. That J implies the existence of a net $\{h_\alpha\}$ of probability densities such that

$$* \begin{cases} f(h_\alpha) - f(h_\alpha * h) \to 0 \\ \text{for all } f \in L_\infty \text{ and all probability densities } h \end{cases}$$

(where, for example,

$$f(h_\alpha) = \int f(g)\, h_\alpha(g)\, d\nu(g))$$

follows from the facts that the probability densities are weak*-dense in the space of all means on L_∞ and that, for f, h_α and h as above, $f(h_\alpha * h) = f * h^*(h_\alpha)$.

WCT ⇒ *SCT*. (The result here is due to M.M. Day, while the proof is due to I. Namioka.)

The assertion $*$ of the previous step may be read as follows: for every probability density h on G, $h_\alpha - h_\alpha * h \to 0$ weakly in L_1 (i.e., for every $f \in L_1^* = L_\infty$, $f(h_\alpha - h_\alpha * h) \to 0$). One then proves that a net of convex combinations of the members of the net $\{h_\alpha\}$ must satisfy SCT. This proof involves the fact that, in a locally convex, linear, topological space (E, τ), the weak and τ-closures of a convex subset coincide. See Greenleaf (1969), Sect. 2.4.

SCT ⇒ *P*. (The result here and its proof are due to A. Hulanicki.)

Let $\varepsilon = 5\varepsilon' > 0$ and compact $K \subset G$ be given, and let k be a fixed probability density on G. By choosing a sufficiently small compact neighbourhood E of the identity, we can have

$$\|k * \psi_E - k\|_1 < \varepsilon'$$

and

$$\|k_g - k\|_1 \quad < \varepsilon', \qquad g \in E,$$

where $\psi_E = I_E/\nu(E)$ (Hewitt and Ross (1963), (20.15) Theorem, p. 293). Since K is compact, we may choose $\{g_1, \dots, g_n\} \subset G$ so that $K \subset \bigcup_1^n g_i E$, and we may assume $g_1 = e$. Putting

$$\psi_i = \psi_{(E g_i^{-1})} = (\psi_E)_{g_i}, \qquad 1 \leqq i \leqq n,$$

we get from SCT a probability density h_α such that

$$\|h_\alpha * \psi_i - h_\alpha\|_1 < \varepsilon', \qquad 1 \leqq i \leqq n,$$

and

$$\|h_\alpha * k - h_\alpha\|_1 < \varepsilon'.$$

Then $h = h_\alpha * k$ is the probability density we need. See Greenleaf (1969), Sect. 3.2, for details.

P ⇒ *S*. We begin by showing the existence of a summing net satisfying (i) of *S* (and then, in Step 5 ahead, we deal with the existence of a summing net satisfying (ii) of *S*, while still satisfying (i) of *S* if *G* is unimodular). This portion of the development must be given in two parts, one part for unimodular groups and one part for non-unimodular groups. The proof for σ-compact, unimodular groups was first given by Chou (1970); it was given independently by Emerson (1974a), who also dealt with σ-compact, non-unimodular groups. We consider first the

Unimodular Case

We first remark that the probability density $h=h(K,\varepsilon)$ of P may be assumed to be a symmetric function, since $h^* * h$ will serve for P as well as h and

$$h^* * h(g^{-1}) = h^* * h(g), \qquad g\in G.$$

We now proceeds by steps, the first two of which are due (except for the part about symmetry) to I. Namioka and C. Ryll-Nardzewski, respectively.

Step 1. Given $\varepsilon>0$, $\delta>0$ and compact $K\subset G$, there exist a compact set $U\subset G$ and a Borel set $N\subset K$ such that $\nu(U)>0$, $\nu(N)<\delta$ and $\nu(Ug\Delta U)/\nu(U)<\varepsilon$ for all $g\in K\backslash N$.

The proof here can proceed like Namioka's (see Greenleaf (1969), Theorem 3.6.3), once we note that all probability densities we get from condition P can be assumed symmetric (by the remark above), as can all simple functions approximating them.

Step 2. Given $\varepsilon>0$ and compact $K\subset G$, there is a symmetric compact set $U\subset G$ such that $\nu(Ug\Delta U)/\nu(U)<\varepsilon$ for all $g\in K$.

Ryll-Nardzewski's idea goes as follows. Assume $\nu(K)>0$ and apply Step 1 to $\varepsilon/2$, $\delta=\nu(K)/2$ and compact set $A=K\cup KK$; the resulting set U will do for Step 2. See Greenleaf (1969), Sect. 3.6 for details.

Step 3. We now assume G is σ-compact and construct a summing sequence for G. Since G is σ-compact, we can write $G=\bigcup_1^\infty O_m$, where each O_m is symmetric and open and has compact closure K_m, $O_m\subset O_{m+1}$, $m=1,2,\ldots$, and $\nu(O_m)\to\infty$. By Step 2, we have, for each m, a symmetric compact set $U_m\subset G$ with $\nu(U_m)>0$ such that

$$\nu(U_m g\Delta U_m)/\nu(U_m)<1/m, \qquad g\in K_m.$$

It follows, by a measure-theoretic argument of Emerson (1968), that $\nu(U_m)\to\infty$. (We note that Chou (1970), and also Emerson and Greenleaf (1968), exhibit ways of choosing the sets $\{U_m\}$ so that it is perhaps a little easier to show $\nu(U_m)\to\infty$.)

We now choose the summing sequence $\{G_n\}$ by induction. Let $G_1=U_1$, and suppose $G_1\subset G_2\subset\ldots\subset G_n$ have been chosen so that

$$G_m=G_m^{-1}\supset O_m, \qquad \nu(G_m g\Delta G_m)/\nu(G_m)<1/m$$

for all $g\in K_m$, $m=1,2,\ldots,n$. Then, choosing $G_{n+1}=U_{j_n}\cup G_n\cup\bar{O}_n$, where

$$j_n\geqq 3(n+1), \qquad \nu(U_{j_n})\geqq 3(n+1)\,\nu(G_n\cup\bar{O}_n)$$

we have

$$\begin{aligned}&\nu(G_{n+1}g\Delta G_{n+1})/\nu(G_{n+1})\\ &\quad\leqq[\nu(U_{j_n}g\Delta U_{j_n})+\nu((G_n\cup\bar{O}_n)g)+\nu(G_n\cup\bar{O}_n)]/\nu(G_{n+1})\\ &\quad\leqq 1/(n+1)\end{aligned}$$

for all $g \in K_{j_n}$, hence for all $g \in K_{n+1}$. Since each compact set $K \subset G$ is contained in K_n for all large enough n, $\{G_n\}$ is indeed a summing sequence as required.

Remark. We note that Emerson (1968) has shown (again via a measure-theoretic argument) that, if $\{G_n\}$ is a sequence of compact subsets of G satisfying $\nu(G_n g \Delta G_n) \to 0$ for all $g \in G$, then this convergence is already uniform (in g) on compact subsets of G. See Sine (1976) for a related result.

Step 4. If G is a locally compact group that is not σ-compact, it can be written as a union of σ-compact open subgroups $G = \bigcup_{\gamma \in J} G^\gamma$; for each γ, $G^\gamma = \bigcup_{n=1}^{\infty} K_n^\gamma$ for suitable compact increasing $\{K_n^\gamma\}_{n=1}^{\infty}$ (as in Step 3). Since each G^γ satisfies (M), we have a summing sequence $\{G_n^\gamma\}_{n=1}^{\infty}$ satisfying

$$\nu(G_n^\gamma g \Delta G_n^\gamma)/\nu(G_n^\gamma) < 1/n, \qquad g \in K_n^\gamma.$$

If we define $(n_1, \gamma_1) \geqq (n_2, \gamma_2)$ to mean

$$G^{\gamma_1} \supset G^{\gamma_2}, \qquad n_1 \geqq n_2, \qquad K_{n_1}^{\gamma_1} \supset K_{n_2}^{\gamma_2} \quad \text{and} \quad G_{n_1}^{\gamma_1} \supset G_{n_2}^{\gamma_2},$$

it follows readily that $\{G_n^\gamma \mid (n, \gamma) \in N \times J\}$ is a net satisfying

$$\nu(G_n^\gamma g \Delta G_n^\gamma)/\nu(G_n^\gamma) \to 0$$

uniformly on compact subsets of G as required.

Non-Unimodular Case

The program here proceeds exactly as in the unimodular case up to the end of Step 2 with the sole exception that the probability densities $h = h(K, \varepsilon)$ of P cannot be assumed symmetric and hence the compact set $U = U(K, \varepsilon)$ of Step 2 satisfying $\nu(U g \Delta U)/\nu(U) < \varepsilon$ for all $g \in K$ cannot be assumed to be symmetric. However, since $\nu(sU) = \Delta(s)^{-1} \nu(U)$, it follows that $V = sU \cup U^{-1} s^{-1}$, which is symmetric, also satisfies

$$\nu(V g \Delta V)/\nu(V) < \varepsilon, \qquad g \in K,$$

if $\Delta(s)^{-1}$ is large enough. Thus we have the conclusion of Step 2 for the non-unimodular case and can apply Steps 3 and 4 to get the desired conclusion.

Step 5. It remains to show that, in the σ-compact case, a summing sequence can be chosen satisfying (ii) of S, while still satisfying (i) of S if G is unimodular. (An argument as in Step 4 then shows how to construct in a non-σ-compact group a summing net for which the analogous assertions hold.) The artful, though basically elementary, arguments alluded to here are almost all due to Emerson (1968); see also Emerson and Greenleaf (1968) and Emerson (1974a).

What needs to be shown is that, given $\varepsilon > 0$ and compact $K \subset G$, we can find a compact (symmetric in the unimodular case) $U \subset G$ such that $\nu(UK)/\nu(U) < 1 + \varepsilon$.

It then follows (using arguments as in Step 3) that we can find compact (symmetric in the unimodular case) sets $\{U_m\}_{m=1}^{\infty}$ with $\nu(U_m K')/\nu(U_m) \to 1$ for all compact $K' \subset G$. (For σ-compact groups, this is the assertion of condition A_1.) In the non-unimodular case, when the sets $\{U_m\}$ are not known to be symmetric, $\{\nu(U_m)\}_{m=1}^{\infty}$ might not tend to infinity; but we can use the non-unimodularity to find a sequence $\{s_m\} \subset G$ for which $\nu(s_m U_m) \to \infty$ and $\nu(s_m U_m K')/\nu(s_m U_m) \to 1$ for all compact $K' \subset G$. Thus we may assume $\nu(U_m) \to \infty$ in both cases. And, to finish, a straightforward argument shows that $\nu(U_m K' \Delta U_m)/\nu(U_m) \to 0$ for all compact $K' \subset G$ and we can use ideas of the latter part of Step 3 to produce the required summing sequence.

So, given $\varepsilon > 0$ and symmetric compact $K \subset G$, we must find a compact $U \subset G$ such that $\nu(UK)/\nu(U) < 1 + \varepsilon$. We know that there is a sequence $\{G_n\}$ of symmetric compact subsets of G with

$$\nu(G_n g \Delta G_n)/\nu(G_n) \to 0, \qquad g \in G.$$

The proof of Emerson (1968) shows how to get the required U in the form $G_{n_0} \setminus E_{n_0}$ for some large n_0, i.e., by chipping away a relatively small amount E_{n_0} from G_{n_0}; in the unimodular case, $U = (G_{n_0} \setminus E_{n_0}) \cap (G_{n_0} \setminus E_{n_0})^{-1}$ also will do and is symmetric. See Emerson (1968), for details.

The only thing we should add here is that a covering property of locally compact groups, which is used in the proof just mentioned and is proved in Emerson and Greenleaf (1967) as a consequence of a difficult, much stronger result in the setting of Lie groups, is now known to have a quite elementary proof. See Milnes and Bondar (1979).

Acknowledgement. The authors wish to thank friends and colleagues who read and commented on the manuscript, notably Lawrence D. Brown, whose advice and encouragement were of great assistance.

References

Berg, C., Christensen, J.P.R.: Sur la norme des opérateurs de convolution. Inventiones math. **23**, 173–178 (1974)

Berk, R.H.: A special structure and equivariant estimation. Ann. Math. Statist. **38**, 1436–1445 (1967)

Berk, R.H.: A remark on almost invariance. Ann. Math. Statist. **41**, 733–735 (1970)

Berk, R.H., Bickel, P.: On invariance and almost invariance. Ann. Math. Statist. **39**, 1573–1576 (1968)

Bondar, J.: On Borel cross-sections and maximal invariants. Ann. Statist. **4**, 866–877 (1976)

Bondar, J.: On a conditional confidence property. Ann. Statist. **5**, 881–891 (1977)

Bondar, J., Milnes, P.: A converse to the Hunt-Stein theorem. Submitted to Ann. Statist (1977)

Brown, L.: A Hunt-Stein theorem (manuscript 1980)

Chou, C.: On topologically invariant means on a locally compact group. Trans. Amer. Math. Soc. **151**, 443–456 (1970)

Chou, C.: The exact cardinality of the set of invariant means on a group. Proc. Amer. Math. Soc. **55**, 103–106 (1976)

Chou, C.: Elementary amenable groups. Illinois J. Math. **24**, 396–407 (1980)

Day, M.M.: Amenable semigroups. Illinois J. Math. **1**, 509–544 (1957)

Day, M.M.: Fixed-point theorems for compact convex sets. Illinois J. Math. **5**, 585–589 (1961) and Correction. Illinois J. Math. **8**, 713 (1964)

Day, M.M.: Amenability and equicontinuity. Studia Math. **31**, 481–494 (1968)

Day, M.M.: Semigroups and amenability. In *Semigroups*, Proceedings of a Symposium at Wayne State University, K.W. Folley (ed.) 5–53. New York-London: Academic Press 1969

Dixmier, J.: Les C^*-algèbres et leurs représentations. Hermann. Paris (1964)
Dunford, N., Schwartz, J.T.: Linear Operators I. New York: Interscience 1958
Eberlein, W.F.: Abstract ergodic theorems and weak almost periodic functions. Trans. Amer. Math. Soc. **67**, 217–240 (1949)
Emerson, W.: Ratio properties in locally compact amenable groups. Trans. Amer. Math. Soc. **133**, 179–204 (1968)
Emerson, W.: Large symmetric sets in amenable groups and the individual ergodic theorem. Amer. J. Math. **96**, 242–247 (1974a)
Emerson, W.: The pointwise ergodic theorem for amenable groups. Amer. J. Math. **96**, 472–487 (1974b)
Emerson, W.: Characterizations of amenable groups. Trans. Amer. Math. Soc. **241**, 183–194 (1978)
Emerson, W.: The Hausdorff paradox for general group actions. J. Functional Anal. **32**, 213–227 (1979)
Emerson, W., Greenleaf, F.: Covering properties and Følner conditions for locally compact groups. Math. Z. **102**, 370–384 (1967)
Eymard, P.: Moyennes Invariantes et Représentations Unitaires. Lecture notes in mathematics no. **300.** Berlin-Heidelberg-New York: Springer 1972
Eymard, P.: Inititation à la théorie des groupes moyennables. In: Analyse harmonique sur les groupes de Lie. (Eymard et al., eds.) 89–107. Lectures notes in mathematics no. 497. Berlin-Heidelberg-New York: Springer 1975
Furstenberg, H.: A Poisson formula for semi-simple Lie groups. Ann. Math. **77**, 335–386 (1963)
Gilbert, J.E.: Convolution operators on $L_p(G)$ and properties of locally compact groups. Pacific J. Math. **24**, 257–268 (1968)
Glasner, S.: Proximal Flows. Lecture notes in mathematics no. **517**. Berlin-Heidelberg-New York: Springer 1976
Greenleaf, F.: Invariant Means on Topological Groups. Van Nostrand math. studies series, no. 16. New York: Van Nostrand 1969
Greenleaf, F.: Amenable actions of locally compact groups. J. Functional Anal. **4**, 295–315 (1969a)
Greenleaf, F.: Ergodic theorems and the construction of summing sequences in amenable locally compact groups. Comm. Pure Appl. Math. **26**, 29–46 (1973)
Greenleaf, F., Emerson, W.: Group structure and the pointwise ergodic theorem for connected amenable groups. Advances in Math. **14**, 153–172 (1974)
Grenander, U.: Probabilities on Algebraic Structures. New York: Wiley 1963
Guivarc'h, Y.: Croissance polynomiale et périodes des fonctions harmoniques. Bull. Soc. math. France. **101**, 333–379 (1973)
Heath, D., Sudderth, W.: On finitely additive priors, coherence and extended admissibility. Ann. Statist. **6**, 333–345 (1978)
Herz, C.: Harmonic synthesis for subgroups. Ann. Inst. Fourier Grenoble. **23**, 91–123 (1973)
Hewitt, E., Ross, K.A.: Abstract Harmonic Analysis I. Berlin-Heidelberg-New York: Springer 1963
Hulanicki, A.: Means and Følner conditions on locally compact groups. Studia Math. **27**, 87–104 (1966)
Iwasawa, K.: On some types of topological groups. Ann. Math. **50**, 507–557 (1949)
Jenkins, J.W.: Growth of connected locally compact groups. J. Functional Anal. **12**, 113–127 (1973)
Johnson, B.E.: Cohomology in Banach Algebras. Amer. Mat. Soc. Memoir no. 127 (1972)
Johnson, B.E.: Perturbations of Banach algebras. Proc. London Math. Soc. (3) **34**, 439–458 (1977)
Kakutani, S.: Two fixed-point theorems concerning bicompact convex sets. Proc. Imp. Acad. Tokyo **14**, 242–245 (1938)
Kesten, H.: Full Banach mean values on countable groups. Math. Scand. **7**, 146–156 (1959)
Kiefer, J.: Invariance, minimax sequential estimation and continuous time processes. Ann. Math. Statist. **28**, 573–601 (1957)
Kiefer, J.: Multivariate optimality results, pp. 255–274 in Multivariate Analysis: Proceedings of the First International Symposium on Multivariate Analysis (Dayton, 1965). P. Krishnaiah (ed.). New York: Academic Press 1966
Lance, C.: On nuclear C^*-algebras. J. Functional Anal. **12**, 157–176 (1973)
Lane, D., Sudderth, W.: Diffuse models for sampling and predictive inference. Ann. Statist. **6**, 1318–1336 (1978)
Lau, A.: Action of topological semigroups, invariant means and fixed points. Studia Math. **43**, 139–156 (1972)

Le Cam, L.: Sufficiency and approximate sufficiency. Ann. Math. Statist. **35**, 1419–1455 (1964)
Le Cam, L.: On the information contained in additional observations. Ann. Statist. **2**, 630–649 (1974)
Lehmann, E.: Optimum invariant tests. Ann. Math. Statist. **30**, 881–884 (1959a)
Lehmann, E.: Testing Statistical Hypotheses. New York: Wiley 1959
Markov, A.: Quelques théorèmes sur les ensembles abéliens. Doklady Akad. Nauk SSSR (N.S.) **10**, 311–314 (1936)
Milnes, P.: Left mean-ergodicity, fixed points and invariant means. J. Math. Anal. Appl. **65**, 32–43 (1978)
Milnes, P., Bondar, J.: A simple proof of a covering property of locally compact groups. Proc. Amer. Math. Soc. **73**, 117–188 (1979)
Nachbin, L.: The Haar Integral. Reprinted edition. New York: Krieger-Huntington 1965
von Neumann, J.: Zur allgemeinen Theorie des Maßes. Fund. Math. **13**, 73–116 (1929)
Paschke, W.L.: Inner amenability and conjugation operators. Proc. Amer. Math. Soc. **71**, 117–118 (1978)
Peisakoff, M.: Transformation Parameters. Thesis. Princeton University 1950
Portnoy, S.: Optimality of best invariant procedures. Mimeographed notes, 26 pp. (c. 1975)
Reiter, H.: Sur la propriété (P_1) et les fonctions de type positif. C.R. Acad. Sci. Paris **258**, 5134–5135 (1964)
Reiter, H.: Classical Harmonic Analysis and Locally Compact Groups. Oxford University Press 1968
Rickert, N.W.: Amenable groups and groups with the fixed point property. Trans. Amer. Math. Soc. **127**, 221–232 (1967)
Rindler, H.: Approximierende Einheiten in Idealen von Gruppenalgebren. Anzeiger der Österreichischen Akademie der Wissenschaften **5**, 37–39 (1976)
Rosenblatt, J.M.: The number of extensions of an invariant mean. Compositio Math. **33**, 147–159 (1976)
Rosenblatt, J.M.: Invariant means on the continuous bounded functions. Trans. Amer. Math. Soc. **236**, 315–324 (1978)
Schwartz, J.: Two finite, non-hyperfinite, non-isomorphic factors. Comm. Pure Appl. Math. **16**, 19–26 (1963)
Sherman, J.: A new characterization of amenable groups. Trans. Amer. Math. Soc. **354**, 365–391 (1979)
Sine, R.: Sequential convergence to invariance in $B(G)$. Proc. Amer. Math. Soc. **55**, 313–317 (1976)
Stein, C.: Approximation of improper prior measures by prior probability measures. Bernoulli-Bayes-Laplace Anniversary Volume (J. Neyman and L. Le Cam eds.) pp. 217–240. Berlin-Heidelberg-New York: Springer 1965
Stewart, J.: Positive definite functions and generalizations: an historical survey. Rocky Mountain Math. J. **6**, 409–434 (1976)
Stone, M.: Necessary and sufficient conditions for convergence in probability to invariant posterior distributions. Ann. Math. Statist. **41**, 1349–1353 (1970)
Stone, M.: Strong inconsistency from uniform priors (with discussion). J. Amer. Statist. Assoc. **71**, 141–125 (1976)
Stone, M., von Randow, R.: Statistically inspired conditions on the group structure of invariant experiments .... Z. Wahrscheinlichkeitstheorie und verw. Geb. **10**, 70–80 (1968)
Templeman, A.A.: Ergodic theorems for general dynamic systems. Dokl. Akad. Nauk SSSR. **176**, 1213–1216 (1967)
Tits, J.: Free subgroups in linear groups. J. of Algebra **20**, 250–270 (1972)
Torgersen, E.N.: Comparison of translation experiments. Ann. Math. Statist. **43**, 1383–1399 (1972)
Wesler, O.: Invariance theory and a modified minimax principle. Ann. Math. Statist. **30**, 1–20 (1959)
Zehnwirth, B.: Invariant least favorable distributions. Macquarie Univ. School of Economic Studies. Tech. Rep. No. 74 (1975)
Zehnwirth, B.: W^* compactness of the class of substatistical decision rules with applications to the generalized Hunt-Stein theorem. Macquarie Univ. School of Economic Studies. Tech. Rept. No. 135 (1977)

Received March 2, 1978; in revised form December 29, 1980